

Universitat de Lleida

Escola Politècnica Superior

Grado en Ingeniería Electrónica Industrial y Automática (GEEIA)

Proyecto de final de grado

IDENTIFICACIÓN DE PAPEL MONEDA MEDIANTE RECONOCIMIENTO DE PATRONES

Autor: Francisco Salcedo Chanza

Director: Francisco Clariá Sancho

Fecha Julio de 2017

Tabla de contenido

1.	Introducción	1
1.1.	<i>Antecedentes</i>	<i>1</i>
1.2.	<i>Motivaciones, Planteamiento del problema</i>	<i>1</i>
1.3.	<i>Objetivos</i>	<i>2</i>
1.4.	<i>Trabajos relacionados.....</i>	<i>2</i>
1.5.	<i>Organización del documento</i>	<i>3</i>
2.	Redes neuronales para el reconocimiento de papel moneda	4
2.1.	<i>Introducción al uso de redes neuronales.....</i>	<i>4</i>
2.1.1.	Aplicaciones actuales	5
2.1.2.	Puntos flacos	5
3.	Procesado de imagen digital.....	6
3.1.	<i>Imagen digital.....</i>	<i>6</i>
3.2.	<i>Procesado en el dominio del espacio</i>	<i>8</i>
3.2.1.	Transformada de nivel de gris.....	8
3.2.2.	Cambio de histograma	10
3.3.	<i>Procesado en el dominio de la transformada</i>	<i>12</i>
3.3.1.	Transformada de Fourier	13
3.3.1.1.	Traslación	14
3.3.1.2.	Rotación	15
3.3.1.3.	Espectro invariante	15
3.3.2.	Reducción en el dominio de Fourier	16
3.4.	<i>Técnicas de detección de contornos.....</i>	<i>16</i>
3.4.1.	Operador de Sobel	17
3.4.2.	Eliminación del <i>background</i> /Substracción de fondo.....	18
4.	Fundamentos matemáticos: PCA e ICA.....	19
4.1.	<i>Introducción al Análisis de Componentes Independientes (ICA)</i>	<i>19</i>
4.1.1.	Ejemplo: La premisa de la fiesta.....	19
4.2.	<i>Separación Ciega de Señales (BBS)</i>	<i>20</i>
4.2.1.	Señales fuente y mezclas de señales	20
4.2.2.	Estrategia general para BSS.....	21
4.2.3.	Matriz de mezcla y matriz recuperadora	22
4.3.	<i>Método ICA</i>	<i>24</i>
4.3.1.	Función Densidad de Probabilidad (<i>pdf</i>)	24
4.3.2.	Entropía con una variable	25
4.3.3.	Independencia del conjunto y distribución marginal.....	27
4.3.4.	Entropía con múltiples variables	27
4.4.	<i>Análisis de Componentes Principales (PCA)</i>	<i>29</i>
4.4.1.	Estrategia de PCA	30

4.4.2.	PCA y señales fuente	30
4.4.3.	PCA como filtro	31
4.5.	Uso de PCA como método de reconocimiento	31
4.6.	Resumen.....	31
5.	Implementación de <i>software</i>.....	33
5.1.	Revisión del grupo de datos	33
5.1.1.	Grupo de datos de entrada	33
5.1.2.	Datos de salida	34
5.1.3.	Consideraciones sobre las muestras	35
5.2.	Métodos implicados: Procesado de imagen	36
5.2.1.	Eliminación de sombras y sustracción de fondo	36
5.2.2.	Detección de contorno.....	37
5.2.3.	Corrección de la inclinación (Giro)	38
5.2.4.	Reducción de la resolución espacial.....	38
5.2.4.1.	Efectos positivos y negativos	39
5.3.	Modelos planteados.....	39
5.3.1.	Clasificación preliminar por dimensión	39
5.3.2.	Modelo idealizado	39
5.3.3.	Modelo de Regiones De Interés (RDI)	40
5.3.4.	Modelo de Bordes	40
5.4.	Aplicación específica de PCA	40
5.4.1.	Columnas/Filas ordenadas	41
5.4.2.	Columnas/Filas acumuladas.....	41
5.4.3.	Reducción de la muestra	41
5.5.	Estudio de la fiabilidad y eficiencia: Selectors.....	41
5.5.1.	Dimensión de la muestra referencia	42
5.5.2.	Discriminación de componentes	42
5.5.3.	Umbral de confianza y tendencia (<i>Threshold and bias</i>)	42
6.	Conclusiones y trabajo a futuro	44
6.1.	Conclusiones.....	44
6.2.	Trabajo a futuro	44
Bibliografía		45
Net y redes		45
Artículos		45
Libros.....		45
Glosario y conceptos.....		46
Índice de figuras		47
Apéndice P: Códigos de programa		2
getIfromdata().....		2

<i>imageProcess()</i>	3
<i>buildpatternV3()</i>	4
<i>rec_patternV3()</i>	5
Apéndice F: Códigos de figuras del capítulo 5	6
<i>Figura 5-2</i>	6
<i>Figura 5-4</i>	6
<i>Figura 5-5</i>	6
<i>Figura 5-6</i>	6
<i>Figura 5-9</i>	6

1. Introducción

“Mejor pan con el corazón feliz que riqueza con exasperación.”

–Amenemopet

1.1. Antecedentes

Los registros en los que se mencionan el uso de “moneda volante” fechan del Siglo VII, en China. Aunque su uso no se hiciera oficial hasta un par de siglos más tarde, era innegable la importancia que tenía como medio de cambio para el comercio. Aún así, no es hasta mediados del siglo XVI (aprox. 1661), que el Banco de Estocolmo comienza a expedir recibos a aquellos depositantes de oro o metales preciosos.¹

Una vez se oficializa el uso de este “papel moneda”, la comodidad que presenta en contra de su predecesor histórico, la moneda metálica, se convierte en la principal razón de su rápida popularización una vez comenzase su uso.

1.2. Motivaciones, Planteamiento del problema

Existen dos razones de gran peso por las que se desarrollan métodos y modelos que reconozcan y clasifiquen los distintos tipos de moneda; la pesadez de realizar un conteo de gran envergadura, con diferentes valores; y la facilidad con la que se puede copiar la imagen de uno de estos. Los bancos con el objetivo de ofrecer un servicio ágil a sus clientes tienen la necesidad de un método automatizado que sustituya la monótona tarea de comprobar y contar billetes o papel moneda. Si bien la tarea no parece tan monótona o repetitiva, cuando el número de billetes y su tipo (refiriéndose al valor monetario) aumentan, las circunstancias cambian.

Debido a la gran cantidad de billetes que se utilizan en la actualidad (no solo cantidad, también en tipo), la necesidad de utilizar maquinas para su conteo aparece, y la necesidad de un procedimiento rápido y fiable para reconocer esta clase de documentos se vuelve indispensable. Dicho proceso no solo se contempla para aquellos casos en que se disponga de un gran número de documentos, sino también para la autenticación de estos. La utilización de imágenes bajo luz infrarroja o ultravioleta muestra distintas imágenes del mismo documento, que aportan posibilidades, tanto a su autenticación como a su clasificación.

Desde el año 2000 se han presentado diversos modelos y soluciones al problema planteado, y la mayoría utilizan un modelo adaptativo conocido como Redes Neuronales (Neural Networks). El método muestra una gran eficiencia, y diversas formas de aplicarse además de distintas entre ellas, cada una con sus inconvenientes y sus ventajas. Aún así, todos estos modelos tienen un aspecto engorroso llamado “fase de aprendizaje” que es inherente al modelo. Principalmente implica un largo periodo de preparación de la maquinaria antes de poder utilizarse. No solo eso, sino que el rendimiento y correcto funcionamiento de la maquinaria pueden depender completamente de esta fase de entrenamiento.

¹ www.riksbank.se/en/The-Riksbank/History/Money-and-power-the-history-of-Sveriges-Riksbank/Stockholms-Banco/.

Lo que se quiere exponer en este proyecto es una solución a este largo periodo de preparación aplicando un método que, si su existencia y aplicación no es novedosa, si lo es la forma y ámbito que se presentan en estas páginas.

1.3. Objetivos

Generales

Utilizando el software MATLAB, estudiar las posibilidades de identificación de imágenes escaneadas de documentos oficiales para su posterior clasificación. Las características esenciales de este tipo de reconocimiento de imágenes es que deben realizarse a alta velocidad y con un alto grado de fiabilidad.

Particulares

- Implementar un algoritmo que permita obtener una imagen limpia del documento, partiendo de una escaneada.
- Diseñar un algoritmo que permita reconocer un documento dado partiendo de una imagen y una base de datos asociada.
- Mejorar los puntos anteriores para realizar su cometido en una cantidad de tiempo menor aligerando la carga de proceso.

1.4. Trabajos relacionados

- ❖ En 2003, Masato Aoba et al. propusieron un método de reconocimiento para Euro en papel moneda utilizando percepción de tres capas y RBF². →Masato Aoba, Tetsuo Kikuchi, Yoshiyasu Takefuji, "Euro banknote recognition system using a three layer perceptron and RBF networks", *IPSI Transaction on Mathematical Modeling and Its Application*, Vol 44, No. SIG 7 (TOM 8), May 2003, Pp. 99-109.
- ❖ En 2003, Ali Ahmadi et al. postularon un método para remover las dependencias no lineales entre variables utilizando SOM³ y extraer las características principales de los datos.→ Ali Ahmadi, Sigeru Omatu, Toshihisa Kosaka, "A Reliable Method for Recognition of Paper Currency by Approach to Local PCA) used local principal component analysis PCA", in *IEEE proceedings of the International Joint Conference on Neural Networks*, 20-24 July 2003, Vol. 2, Pp. 1258 – 1262.
- ❖ En 2008, D. A. K. S. Gunaratna et al. propusieron el sistema "SLCRec" con una función transformación lineal especial que se adapta para cancelar patrones de ruido de trasfondo sin afectar las imágenes características de billetes y reparar imágenes de interés. →D. A. K. S. Gunaratna, N. D. Kodikara and H. L. Premaratne, "ANN Based Currency Recognition System using Compressed Gray Scale and Application for Sri Lankan Currency Notes-SLCRec", in *proceedings of world academy of science, engineering and technology*, Nov 2008, vol. 35, ISSN 2070-3740, Pp. 235-240.
- ❖ En 2010, Kalyan Kumar et al. presentaron un sistema de reconocimiento para billetes utilizando redes neurales ensambladas (ENN⁴), particularmente para TAKA (moneda de Bangladeshi). Para probar la eficacia del método los autores lo compararon con otros métodos como Modelo Oculto de Markov (HMM⁵), RBF, y SLCRec.→Kalyan Kumar Debnath,

² Radial Basis Function.

³ Self-Organized Map.

⁴ Ensemble Neural Network.

⁵ Hidden Markov Model.

Sultan Uddin Ahmed, Md. Shahjahan, "A Paper Currency Recognition System Using Negatively Correlated Neural Network Ensemble", *Journal Of Multimedia*, December 2010, Vol. 5, No. 6, Pp. 560-567.

- ❖ En 2010, Junfang Guo et al. utilizaron un algoritmo "block-LBP" para la extracción característica de papel moneda. Block-LBP es una versión mejorada del tradicional método patrón binario local (LBP⁶) → Junfang Guo, Yanyun Zhao, Anni Cai, "A reliable method for paper currency recognition based on LBP" in *proceeding of 2nd IEEE International Conference on Network Infrastructure and Digital Content*, 24-26 Sept. 2010, Pp. 359 – 363.
- ❖ En 2011, Hai-dong et al. propusieron un método numérico de reconocimiento de papel moneda que ordenaba los valores propios de un vector, y hallar el mejor umbral y tendencia. → Hai-dong Wang, Leye Gu, Linping Du, "A paper currency number recognition based on fast Adaboost training algorithm", in *IEEE proceedings of International Conference on Multimedia Technology (ICMT)*, Pp. 4772 – 4775.
- ❖ En 2012, Chetan B. V. et al. propusieron un método de reconocimiento de papel moneda de dos fases basado en la coincidencia entre la imagen de un billete con una base de datos de imágenes del mismo tipo de billete. → Chetan B. V., Dr. P. A. Vijaya, "A Robust Side Invariant Technique of Indian Paper Currency Recognition", *International Journal of Engineering Research & Technology (IJERT)*, May – 2012, Vol. 1 Issue 3, Pp. 1-7.

1.5. Organización del documento

El proyecto se organiza en 6 capítulos y 4 Apéndices con la siguiente estructura:

- Capítulo 2: presenta una explicación del objeto entorno al que trata el proyecto y revisa, de forma superficial, la tecnología empleada actualmente para resolver el problema que plantea.
- Capítulo 3: lista y muestra los aspectos en relación al tratamiento de imágenes digitales que se han utilizado y revisado para el desarrollo del proyecto.
- Capítulo 4: explica los fundamentos matemáticos en los que se apoya el proyecto así como dar una idea generalizada de su ámbito de aplicación así como una defensa de la decisión tomada al escoger el uso del modelo de PCA.
- Capítulo 5: se describen los procedimientos, concesiones y demás resultados que se han llevado a cabo tras la aplicación práctica de los conocimientos presentados en los capítulos 2, 3, y 4. También se describen de forma genérica las conclusiones derivadas de cada experimentación, test y prueba sobre los objetos del proyecto.
- Capítulo 6: presenta las conclusiones a las se obtuvieron a raíz del desarrollo del proyecto y sus resultados.
- Apéndice F: Apéndice donde se pueden encontrar los códigos de entorno MatLab utilizados para la construcción de algunas de las figuras presentes en el documento.
- Apéndice P: Apéndice donde figura una lista ordenada con los principales códigos de entorno MatLab utilizados durante el proyecto.

⁶ Local Binary Pattern.

2. Redes neuronales para el reconocimiento de papel moneda

Los procedimientos automatizados para el reconocimiento de billetes surgen de la necesidad de realizar esta tarea con rapidez. La importancia de este documento radica en su fácil transporte y manejo, y no es extraño ver grandes cantidades de estos juntos.

Uno de los puntos importantes del billete moderno es la complejidad de su imagen. Actualmente, cada billete tiene su propio “rostro”, que difiere claramente de aquellos de distinto valor aunque misma moneda. Esta complejidad es a su vez un modo de autenticidad y un método de clasificación; la imagen impresa sobre ellos es fácilmente reconocible a ojos humanos y es fácil decir a cuanto equivalen. A su vez, diferentes marcas e impresiones que se ven a distintas radiaciones permiten asegurar la autenticidad de estos trazos de papel.

Como se ha dicho, estas imágenes son complejas, incluso para la vista humana. Entonces, para el reconocimiento de estos billetes, se busca un método que pueda equipararse a la percepción visual del ojo humano y tenga la potencia y rapidez de una computadora. Las redes neuronales aportan ambos rasgos.

Las redes neuronales son un modelo del que su desarrollo se ha visto afectado por las limitaciones de la tecnología de la época, pero aún así en las últimas décadas ha demostrado tener un gran potencial en muchos campos distintos. En el reconocimiento de billetes su potencial ha causado que diversidad de modelos surjan, principalmente basándose en la interpretación del billete como un conjunto de patrones.

2.1. Introducción al uso de redes neuronales

Una red neuronal es un sistema biológico formado por un gran conjunto de células nerviosas (neuronas) unidas entre ellas. El cerebro y gran parte del sistema nervioso está formado por ellas, y son responsables del aprendizaje y de la interpretación de aquello que nos rodea.

En el intento de recrear esta capacidad de adaptación y desarrollo, en la década de los años cuarenta (1940) se definen unos nuevos modelos teóricos de funciones mentales y actividad cerebral con la finalidad de desarrollar Redes Neuronales Artificiales⁷. Se esperaba obtener sistemas adaptativos, capaces de modificar su funcionamiento y enlaces con el objetivo de alterar la información contenida en ellos. En esencia, estos sistemas tenían que ser capaces de desarrollarse a sí mismos para cumplir su función.

Su desarrollo e investigación pasan desapercibidos durante los cincuenta y sesenta, y se llega a un bloqueo que detiene su progreso a mitad de los 60. Este receso se debe principalmente a las deficiencias de la tecnología de computación en ese momento; los sistemas que se presentaban para Redes Neuronales no contaban con un procesamiento en paralelo y la gestión se realizaba de forma secuencial en su totalidad. El diseño era mucho más rígido; las variables que intervenían en la función se definían durante el proceso de aprendizaje y se mantenían constantes. Con el desarrollo de la tecnología y un mayor entendimiento del cerebro, se implementaron sistemas de múltiples entradas, y las tareas de cómputo se realizaban a la vez que estas modificaban las funciones internas en un ciclo de retroalimentación.

⁷ *Artificial Neural Networks (ANN).*

El interés por las Redes Neuronales sucede por la aplicación para el reconocimiento de patrones⁸ en entornos reales, discriminación no lineal, y la resolución de problemas que los métodos de computación convencionales no pueden manejar.

Las Redes Neuronales Artificiales (Artificial Neural Networks, ANN) se han desarrollado como disciplina tecnológica que puede desarrollar automáticamente capacidades operativas para adaptar su respuesta al entorno. La evolución de las redes neuronales artificiales no ha sido fácil. La investigación en este campo ha pasado desapercibida desde los cincuenta (1950). Es el periodo de oro de las redes neuronales artificiales queda entre los años cincuenta y los sesenta (1950-60). A mitad de los años sesenta, el primer éxito de la neurocomputación llevó a una parada por algunas razones.

Durante estos años tranquilos (de 1967 a 1982) poca investigación de neurocomputación se ha llevado a cabo de forma explícita en los Estados Unidos. Un gran hito de la investigación de redes neuronales llegó bajo el nombre de reconocimiento de patrones, modelado biológico, y procesado de señal adaptativo. En 1983 el entusiasmo por las redes neuronales reapareció.

2.1.1. Aplicaciones actuales

En la actualidad, la redes neuronales no pertenecen un campo específico de forma exclusiva, y pueden hallarse en gran variedad de aplicaciones: Cancelación adaptativa de ruido en telecomunicaciones; evaluador de mortalidad; reconocimiento de lenguaje; marketing para aerolíneas en la venta de billetes y distribución de asientos; etc.

En ciencias y tecnologías del entorno: predicción climatológica, análisis de tendencias, lectura de escritura a mano, asesoramiento en riesgo financiero, control de procesos de producción, control de calidad.

2.1.2. Puntos flacos

La flaqueza de las redes neuronales es, irónicamente el punto que caracteriza el método, la fase de aprendizaje. La fase de aprendizaje es la etapa que más influye en los resultados una vez el método está listo para utilizarse. Para la fase de aprendizaje se utilizan unos prototipos de los patrones que deben clasificarse más tarde. Estos prototipos son muy numerosos, e incluso deben seleccionarse con cautela para que el modelo no converja en una única solución o un conjunto de pocas soluciones posibles.

Un gran número de prototipos implica un gran periodo de tiempo dedicado exclusivamente a la fase de aprendizaje. Esto hace que una vez la fase de aprendizaje ha terminado su utilidad queda definida, y que el modelo generado no pueda recuperarse.

⁸ *Pattern recognition.*

3. Procesado de imagen digital

Una imagen digital es, en esencia, un gran número de datos organizados en píxeles que, de forma individual, no aportan ninguna información. Es cuando se analizan en conjunto cuando la información contenida en ellos gana importancia y significado. Uno podría preguntarse entonces, como determinar la importancia de cada uno de estos píxeles desde un punto de vista digital.

El procesado de una imagen digital engloba tanto la conversión de una imagen a datos tratables como su “preprocesado”, que ayuda no solo a comprender el sentido global sino a determinar los pasos que se llevarán más tarde para su conversión en los datos que se desean.

El preprocesado es un paso importante cuando se trabaja con datos de gran envergadura⁹ como lo es una imagen. Uno de los aspectos importantes de este proceso es que no altera la información original de la imagen, sino la forma de visualizarla y de entenderla. Este paso, en ocasiones tan simple como aplicar un filtro, puede ahorrar gran cantidad de tiempo cuando se extraen estos datos. Uno de los inconvenientes es la variedad de parámetros relacionados, que hacen difícil (sino imposible) determinar una solución generalizada.

En el caso planteado, el preprocesado de imágenes se utilizará para localizar el conjunto de puntos relevantes dentro de la imagen y descartar el resto. Estos puntos relevantes serán los pertenecientes al conjunto que forma la imagen del billete en la captura.

3.1. Imagen digital

Primeramente consideremos la definición de la palabra imagen como: “La figura de una persona o cosa captada por el ojo, por un espejo, un aparato óptico, una placa fotográfica, etc., gracias a los rayos de luz que recibe y proyecta. “

Entendamos pues, que la imagen digital es aquella imagen que ha sido digitalizada; Es decir, que la información contenida en ella ha sido convertida según un método específico para poder reinterpretarse más tarde en un entorno digital (por ejemplo, una pantalla o monitor). Se debe hacer hincapié en un aspecto que suele pasar desapercibido para quien lo desconoce; una imagen una vez convertida no puede reproducirse al formato físico y obtener la misma imagen original. Ejemplo: se digitaliza una fotografía antigua para su preservación y/o reproducción. Por ello, la digitalización es un proceso en el que se pierde información, y es irreversible.

Las imágenes digitales (de ahora en adelante referidas como imagen para abreviar) disponen de limitaciones establecidas por su método de conversión. Específicamente, se explica el aspecto relacionado con el color. El ojo humano capta la luz de forma analógica; para expresarlo sencillamente, sin una clara diferenciación entre dos magnitudes muy próximas. La conversión digital, que está limitada por la memoria virtual disponible, muestrea la señal de luz con cierta periodicidad (ejemplo: 1píxel/mm), que se convierten en dos parámetros para la imagen digital; el nivel de color y la resolución de imagen, respectivamente.

⁹ Anglismo: *Large Data Sets*.

El punto importante sobre el nivel de color es que, aunque no necesariamente es igual para todos los formatos de imagen, se establece que cualquier imagen digitalizada se interpreta en una base de 256 niveles de color (con un valor posible entre 0 y 255, ambos incluidos). Esta base aumenta en según qué formatos, pero la mayoría de aparatos que reproducen imágenes tienen preestablecido que la información es en base a este estándar de 256 niveles por color¹⁰. Tomando ventaja de la fisiología del ojo humano, solo se toman en cuenta las intensidades de tres colores; rojo, verde y azul. Esta forma de digitalizar corresponde al formato RGB-8, y sus resultados como efectos se intentan mostrar en la Figura 3-1, y Figura 3-2 presentadas a continuación.



Figura 3-1, Recreación de un cartel holográfico "Tranquillity's STOP".

La resolución determina, de forma general, el nivel de detalle que dispone la imagen. A mayor resolución, mejor definición de formas, y se pueden apreciar elementos más pequeños. Se expresa en píxeles en ancho y alto (ejemplo 1920x1080). La resolución representa la distancia real entre dos píxeles de la imagen y determinará su tamaño. Matemáticamente, puede entenderse como una función de dos variables x e y (con x e y como ancho y largo), o como una matriz con número de columnas igual a su ancho y número de filas igual a su alto. De modo que cada elemento de dicha matriz corresponde a cada píxel de la imagen.

¹⁰ Este formato de digitalización corresponde al formato RGB-8. El 8 representa el número de bits utilizados para cada color.

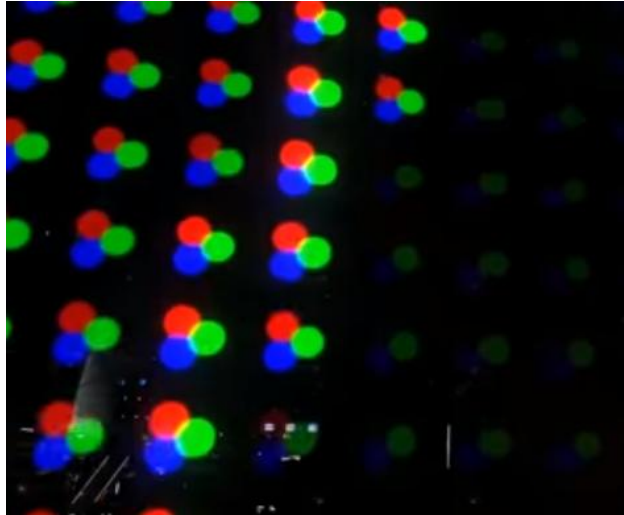


Figura 3-2, Detalle de los pixeles del cartel.

La importancia de esta explicación reside en que estos elementos (resolución y nivel de color) afectan a la cantidad de datos que forman la imagen y a la carga de procesamiento para trabajar con ella.

Uno de los objetivos de este proyecto es alcanzar una solución para el problema planteado que aporte velocidad a la tarea. Esto se traduce como “aligerar la carga de procesamiento”. Por ello, no solo se trabajará con imágenes en escala de grises, sino que en algún punto (además de por necesidad) también se reducirá la resolución de la imagen.

3.2. Procesado en el dominio del espacio

Cuando nos referimos al dominio del espacio al tratar con una imagen, nos referimos a que se trabaja sobre cada uno de sus pixeles. Concretamente, en el valor asociado a ellos.

Trabajar sobre el dominio del espacio permite visualizar el contenido de una imagen a medida que estos se realizan. Con estas técnicas se ejerce un control más directo sobre la imagen, que se traduce en un mayor efecto visual de los cambios que se están llevando a cabo y una menor capacidad de adaptación a valores o casos inesperados.

Habitualmente se trabajan sobre la imagen en pequeñas secciones de esta y no de punto en punto. Estas secciones, llamadas vecindario (del punto), se forman a través del pixel que se pretende tratar y aquellos que lo rodean, formando conjuntos de 3x3 o 5x5 habitualmente. Esto resulta en una carga de cómputo proporcional al número de pixeles incluidos. Sin embargo, permite un mayor conocimiento contextual dentro de la imagen que no se obtiene tratando punto por punto. Podría decirse que, el uso de estos puntos vecinos aporta una mayor comprensión del entorno del pixel.

3.2.1. Transformada de nivel de gris

Como se menciona en el apartado 3.1 (Imagen digital) y posteriormente en el apartado 5.1 (Revisión del grupo de datos), las imágenes que conforman la base de datos para el proyecto se encuentran en un formato monocromático de grises. Principalmente, implica que en la información contenida en la imagen no hay distinción de color. La falta de esta información no implica necesariamente un inconveniente; según el caso en que se encuentra

es incluso favorable disponer de una imagen monocromática. Esta es más ligera en cuanto a peso de memoria¹¹ y por ello más fácil de tratar.

El objetivo del preprocesado de imágenes es intentar mejorar su visualización; una posibilidad es ajustar los niveles de gris que conforman la imagen. Con esta mejora, se puede conseguir que una imagen que apenas tiene ninguna diferencia entre sus distintos niveles de gris, y por ello su contenido es indistinguible, conforme una imagen clara.



Figura 3-3, Fotografía de un edificio acristalado.



Figura 3-4, Mejora aplicando un escalado lineal.

¹¹ De forma teórica, una imagen RGB ocupa como mínimo el triple de espacio en memoria que una monocromática o escalada en grises. El valor exacto depende de su formato y compresión.

Sin embargo, el objetivo de este proyecto no es mejorar la vista de las imágenes (no de la imagen original al menos) con el objetivo de visualizarla, y la aplicación de estos métodos no se estudiarán con esta intención, sino con la de mejorar su estado para algoritmos posteriores.

La aplicación de una transformada en nivel de gris afecta a toda la imagen por igual y se representa como una tabla de conversión o una función algebraica. Esta transformación altera el valor original del pixel por otro en función de este y no de su posición dentro de la imagen. En la Figura 3-5 se muestran algunas representaciones de transformaciones de nivel de gris.

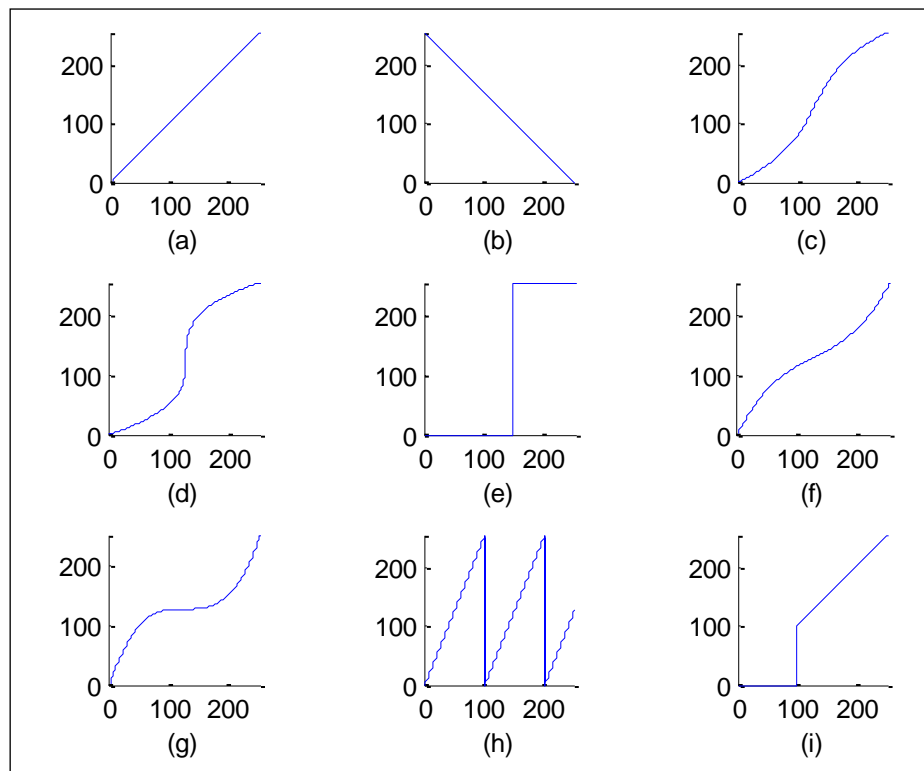


Figura 3-5, Diferentes aplicaciones de transformada de nivel.

Aquellas con expresión simple como (a) y (b), se aplican para alterar la diferencia de intensidad entre distintos niveles; haciendo tonos oscuros más oscuros, tonos claros más claros, etc. Otras transformaciones podrían tener como objetivo aumentar el contraste en tonos intermedios (c), (d), (f) y (g), filtrar un rango de tonalidades (i), o convertir la imagen en binaria (e) a cierto valor.

3.2.2. Cambio de histograma

Considerando una imagen desde un punto de vista matemático, esta se puede interpretar como una matriz M con w columnas y h filas (donde estas corresponden a la resolución de la imagen), y sus elementos responden a la expresión $M = f(x, y)$, donde f es la función particular que recrea la imagen. f es una función definida en para los valores positivos de x e y , y esta acotada a valores enteros positivos entre 0 y 255. Diremos que para cada pareja de valores de x e y la función f tiene como resultado un valor *racotado* por los límites de f . Si bien el tecnicismo puede resultar confuso, es necesaria su aclaración para dar sentido al estudio de la imagen a través de su histograma.

$$M_{h,w} = f(x, y)$$

Ecuación 3.1

$$f: \mathbb{R}^2 \rightarrow \mathbb{R} = [0, 255]$$

Ecuación 3.2

Por definición, el histograma de una imagen muestra la frecuencia relativa de un mismo valor en sus elementos. Cuando se trata de una imagen en escala de grises, se entiende que cada valor se corresponde con un mismo nivel de color. Para una visualización cómoda de su contenido, en un histograma se representan las repeticiones para determinados intervalos de valores que no suelen corresponderse con los niveles individuales de la variable representada. La Figura 3-6 intenta reflejar la relación entre el histograma de una variable (su frecuencia relativa) y la grafica de la densidad de probabilidad.

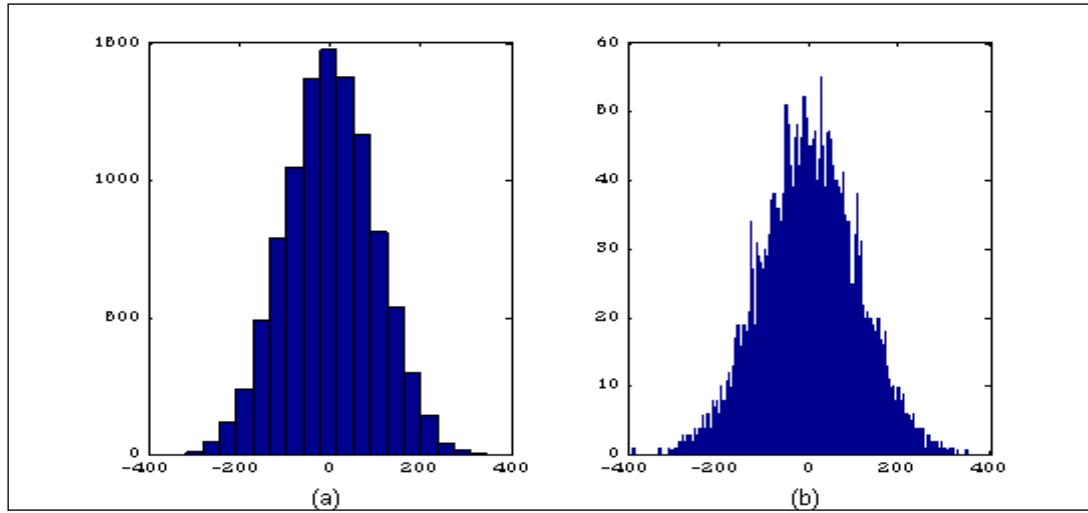


Figura 3-6, Relación entre el histograma de una variable y grafica densidad de probabilidad.

La modificación de histograma se utiliza en mejoras de contraste selectivo. Utilizando una función transformada, se altera la forma de su *pdf* hasta obtener una grafica que se adapte mejor a las necesidades del problema. La Ecuación 3.3 y la Ecuación 3.4 muestran como se relacionan el dominio r y el dominio s .

$$s = T(r); r_0 \leq r \leq r_j$$

Ecuación 3.3

$$r = T^{-1}(s); s_0 \leq s \leq s_k$$

Ecuación 3.4

La expresión de la transformada puede ser en forma de tabla de valores o como la expresión matemática de una curva. Sin importar como se exprese, la función de la transformada deberá satisfacer las siguientes condiciones para ser válida:

1. Debe ser una función univariable¹² y monótona creciente¹³ para evitar situaciones ambiguas en el intervalo de $0 \leq r \leq 1$ tras normalizar.
2. s , que es $T(r)$, debe hallarse en el intervalo entre 0 y 1 para $0 \leq r \leq 1$.

3.3. Procesado en el dominio de la transformada

El procesado de imágenes en el dominio de la transformada surge de la necesidad de tratar con grandes conjuntos de datos. En el dominio de la transformada, algunas de las características propias de una imagen son resaltadas a partir de su espectro. Además, este nuevo dominio ofrece nuevas propiedades de cómputo que podrían llegar a distintos resultados en comparación con el procesado en el espacio visto en el apartado anterior, Procesado en el dominio del espacio.

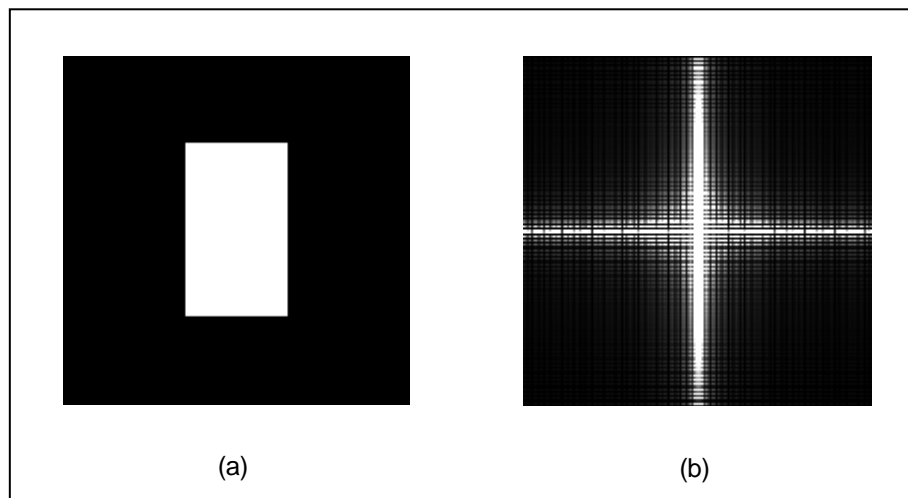


Figura 3-7, (a) Imagen artificial rectangular y (b) su espectro.

El primer punto del procesado es su transformada y la interpretación de esta. Para expresarlo de forma simple, en el dominio de la transformada una misma imagen no es comprensible a simple vista, como se muestra en la Figura 3-7. Siendo el caso, es necesario encontrar una forma de convertir la imagen $f(x,y)$ al dominio de la transformada. Se buscará pues, una función invertible g que permita convertir el dominio de $f(x,y)$ al dominio de la transformada. Esto supone un aumento de etapas dentro del proceso general, tal como se intenta expresar en el diagrama de la Figura 3-8; una vez los cambios se han hecho en el dominio de la transformada, estos deben invertirse al dominio original para poder apreciarse. Este motivo hace que trabajar en el dominio de la transformada pueda ser en algunos casos infructuoso; el nuevo dominio no facilita suficiente la tarea y la carga de cómputo aumenta hasta niveles innecesarios.

¹² Que solo es dependiente de una variable. Ejemplo: $f(x)$

¹³ Véase en el glosario.

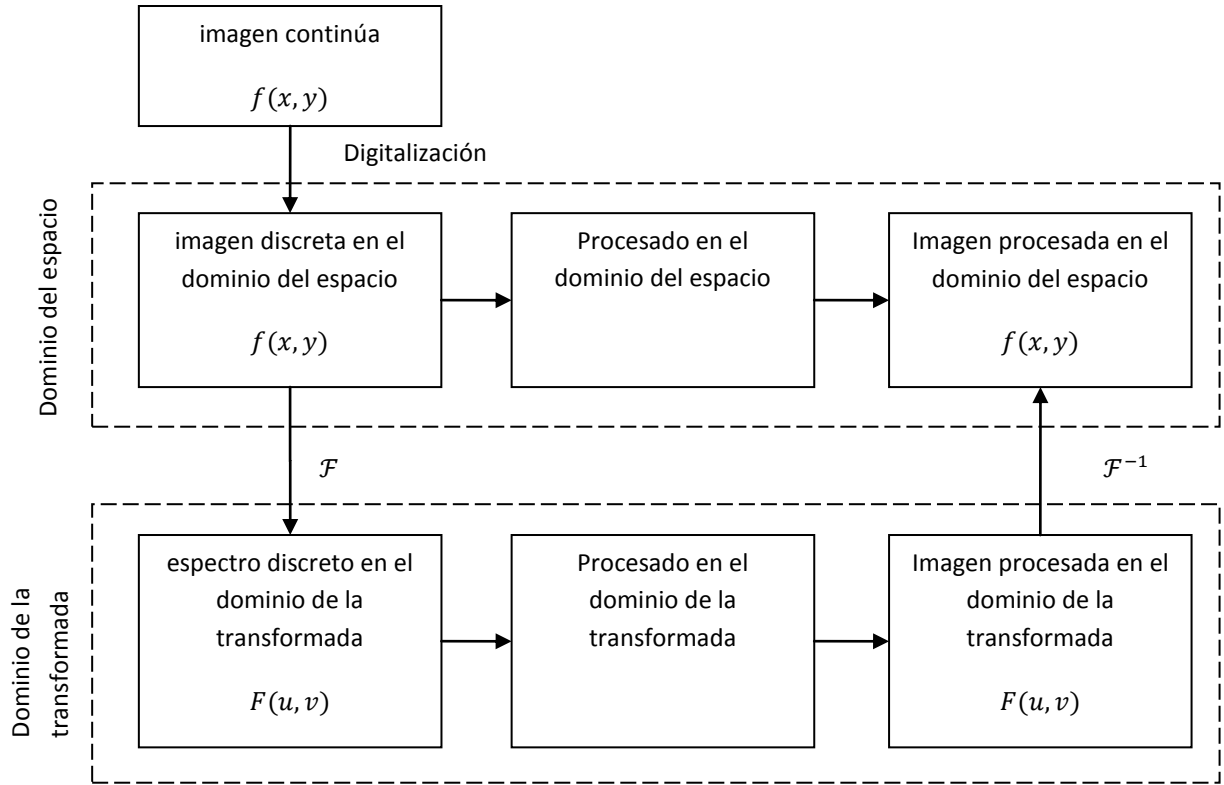


Figura 3-8, Esquema de las etapas de la imagen según el dominio utilizado.

Seguidamente, se expone el caso particular de la transformada de Fourier, tal vez la más sencilla de comprender y visualizar debido a sus propiedades en el nuevo dominio.

3.3.1. Transformada de Fourier

El dominio de Fourier corresponde al dominio de la frecuencia, y la función de transformada \mathcal{F} para un dominio continuo (como es el tiempo t) se muestra en la Ecuación 3.5 y Ecuación 3.6. Sin embargo para el caso de imágenes, se dispone de un dominio discreto de dos dimensiones, que se expresa como la Ecuación 3.7 y Ecuación 3.8.

$$F(j\omega) = \int_{-\infty}^{\infty} e^{-j\omega t} f(t) dt \quad \text{Ecuación 3.5}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} F(j\omega) d\omega \quad \text{Ecuación 3.6}$$

$$F(u, v) = \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-\frac{j2\pi(ux+vy)}{N}} \quad \text{Ecuación 3.7}$$

$$u, v = 0, 1, 2, \dots, N-1$$

$$f(x, y) = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) e^{\frac{j2\pi(ux+vy)}{N}} \quad \text{Ecuación 3.8}$$

$$x, y = 0, 1, 2, \dots, N-1$$

El dominio de Fourier, como se menciona anteriormente en este mismo capítulo, se corresponde al dominio de la frecuencia. En este dominio se ve reflejada sobre todo, la información sobre los patrones y la geometría en la imagen.

3.3.1.1. Traslación

En ocasiones puede suceder que la visualización del dominio de Fourier sea de difícil interpretación porque sus componentes destacables se concentren en un borde o en una esquina.

Supongamos $F(u, v)$ la función transformada de $f(x, y)$. Si se multiplica $f(x, y)$ por un factor exponencial $\exp(j2\pi(u_0x + v_0y)/N)$, y se realiza la transformada del producto, el origen del espectro de frecuencias será (u_0, v_0) , y su expresión de su transformada es $F(u - u_0, v - v_0)$. Se deduce de la Ecuación 3.10 que la forma de este nuevo espectro coincide con la original, tan solo que desplazada en sus ejes u y v . Esta propiedad puede aplicarse también en el dominio del espacio cuando se multiplica su espectro por $\exp(-j2\pi(ux_0 + vy_0)/N)$ con resultados equivalentes. Para el caso planteado a principio del párrafo, se propondría desplazar u_0 y v_0 a $N/2$.

$$f(x, y) \xrightarrow{\mathcal{F}} F(u, v) \quad \text{Ecuación 3.9}$$

$$f(x, y) \cdot \exp\left(\frac{j2\pi(u_0x + v_0y)}{N}\right) \xrightarrow{\mathcal{F}} F(u - u_0, v - v_0) \quad \text{Ecuación 3.10}$$

$$F(u, v) \cdot \exp\left(\frac{-j2\pi(ux_0 + vy_0)}{N}\right) \xrightarrow{\mathcal{F}^{-1}} f(x - x_0, y - y_0) \quad \text{Ecuación 3.11}$$

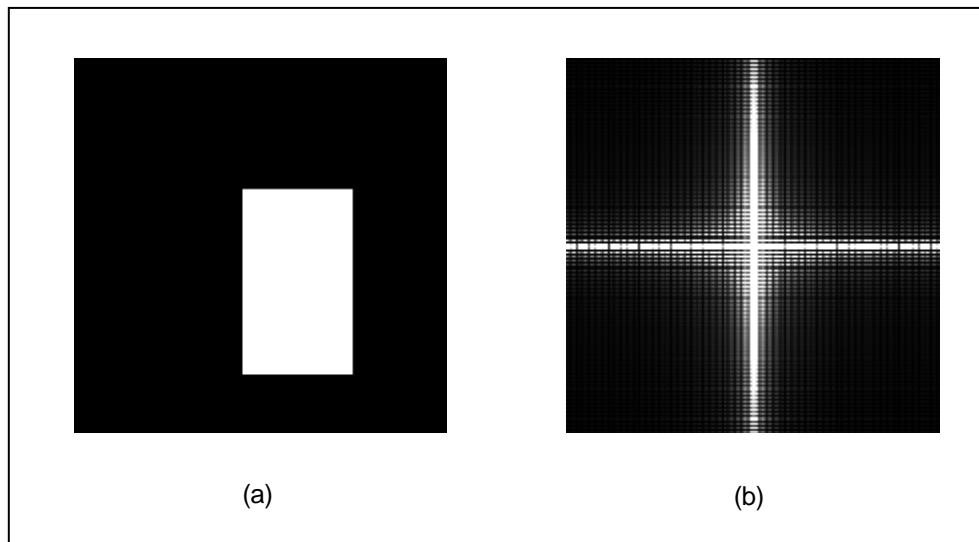


Figura 3-9, (a) Imagen desplazada y (b) su espectro tras el desplazamiento.

Se desea matizar que, pese a la transformación llevada a cabo, las magnitudes en el dominio de Fourier no se alteran aparte de la traslación. Es decir que los factores exponenciales

$\exp(-j2\pi(ux_0 + vy_0)/N)$ y $\exp(j2\pi(u_0x + v_0y)/N)$ no alteran el modulo de las funciones a las que multiplican.

3.3.1.2. Rotación

Si una imagen en el dominio del espacio es rotada un ángulo $\Delta\theta$, su transformada varía un ángulo $\Delta\phi$. Es decir, expresada una imagen y su transformada de Fourier en coordenadas polares $f(\rho, \theta + \Delta\theta)$ su transformada es $F(r, \phi + \Delta\phi)$, donde $\Delta\theta$ y $\Delta\phi$ son iguales.

$$f(\rho, \theta + \Delta\theta) \xrightarrow{\mathcal{F}} F(r, \phi + \Delta\phi); \Delta\theta = \Delta\phi$$

Ecuación 3.12

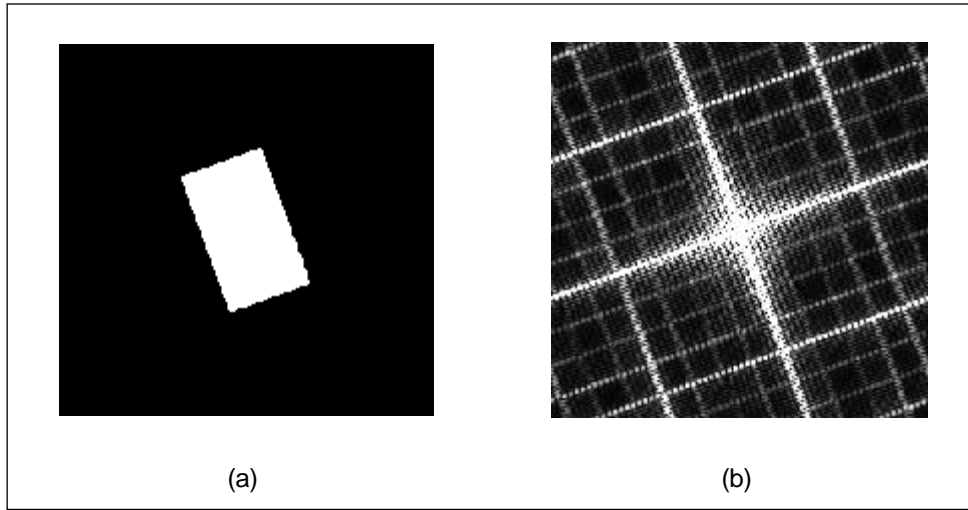


Figura 3-10, (a) Imagen artificial rotada y (b) su espectro tras rotar.

3.3.1.3. Espectro invariante

La propiedad más interesante que muestra la transformada es tal vez su “invariabilidad” ante la posición de una geometría en el dominio del espacio. Esto se entiende de la siguiente forma; si suponemos una imagen *vacía* con un único objeto, su espectro, que se corresponde con sus componentes en frecuencia, es igual independientemente de donde se localiza dicho objeto.

Esta propiedad puede incluso discutirse a través de la traslación, vista anteriormente. Recordando la Ecuación 3.11 sobre la traslación en el dominio del espacio, se expresa su función transformada de Fourier como una parte real y otra imaginaria. En la gráfica, solo se ven reflejada la parte real, que se forma a partir del modulo. Así pues, el modulo de la transformada, que se corresponde a la intensidad de los puntos (u, v) , se expresa según la Ecuación 3.13, y el modulo de $\exp(-j2\pi(ux_0 + vy_0)/N)$ debe ser unitario para corresponderse con la propiedad de translación.

$$|F(u, v)| = \left| F(u, v) \cdot \exp\left(-\frac{j2\pi(ux_0 + vy_0)}{N}\right) \right|$$

Ecuación 3.13

3.3.2. Reducción en el dominio de Fourier

En dependencia de su contenido, toda imagen tiene un espectro que le corresponde. De forma reiterada se ha mostrado como la información contenida en el espectro se encuentra concentrada en el centro de la imagen tras desplazar las coordenadas cero (u_0, v_0) a la muestra $N/2$.

A la luz de este hecho, se hace interesante saber que si la mayoría de información contenida en el espectro se concentra cerca del centro, cuanta de esta es realmente necesaria para una reconstrucción aceptable de la imagen. Esta medida es dependiente de la propia imagen, y sin un análisis de esta, no es una valoración predecible. Aún así, si se está dispuesto a sacrificar la definición de la imagen, descartar el espectro más alejado del centro incurrirá en una reducción de la carga de proceso.

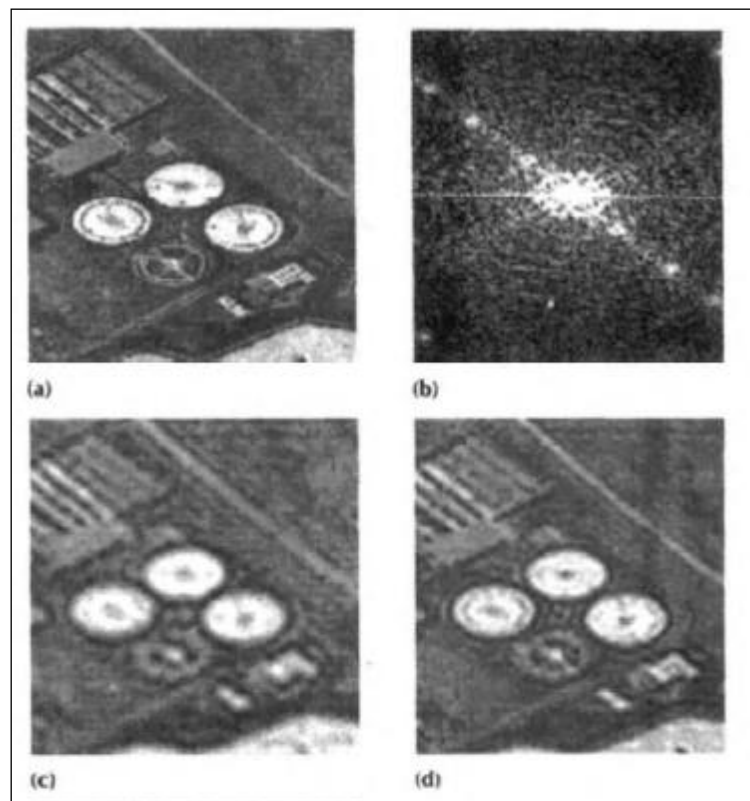


Figura 3-11, Efecto de la reducción del espectro: (a) Imagen original, (b) Interpretación de su espectro, (c) Imagen restaurada de un 5% del espectro, (d) Imagen restaurada de un 10% de su espectro.¹⁴

3.4. Técnicas de detección de contornos

La diferencia entre las técnicas de detección de contornos y la detección de patrones puede resumirse en el conocimiento del “objeto” que se espera encontrar. Habitualmente las técnicas de detección de patrones se apoyan en una base de datos con distintos elementos que se corresponden a patrones de objetos, colores, etc. Estos se utilizan para analizar la imagen y así identificar los distintos objetos que lo forman. La detección de contornos está encarada en detectar importantes cambios de nivel en la imagen, en otras palabras “saltos” de

¹⁴ Figura extraída de Sing-Tze Bow, *Pattern Recognition and Image Preprocessing Second Edition, Revised and expanded*; ISBN:0-8247-0659-5.

intensidad o color. Con la información de estos saltos de nivel, pueden delimitarse las regiones de la imagen y clasificarse como objetos concretos. Así pues, en comparación con la detección de formas y patrones, la detección de contornos presenta una mayor flexibilidad con la distinción de objetos y sus formas. Por el contrario, carece de la agilidad que puedan aportar los otros métodos.

Seguidamente se presentan unos métodos relacionados con la detección de contornos y bordes. Se recuerda que la detección de contornos es un proceso intermedio; tiene como fin delimitar una región de la imagen, aunque puede también aplicarse para la detección de formas sin utilizar la transformada de Fourier.

3.4.1. Operador de Sobel

El método de Sobel consiste en la aplicación de un operador direccional a la imagen en cuestión. Este operador calcula un gradiente diferencial de la intensidad de la imagen, donde el resultado es modulo, que refleja la intensidad del borde.

$$S_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$$

Ecuación 3.14

$$S_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

Ecuación 3.15

Los operadores de Sobel, también tratados como filtros de Sobel, son elementos asociados a una dirección concreta. Es decir, la aplicación del operador S_x tendrá un diferente resultado que si se aplica el operador S_y . No solo eso, sino que también serán distintos los resultados entre S_x y su operador opuesto S_{-x} .

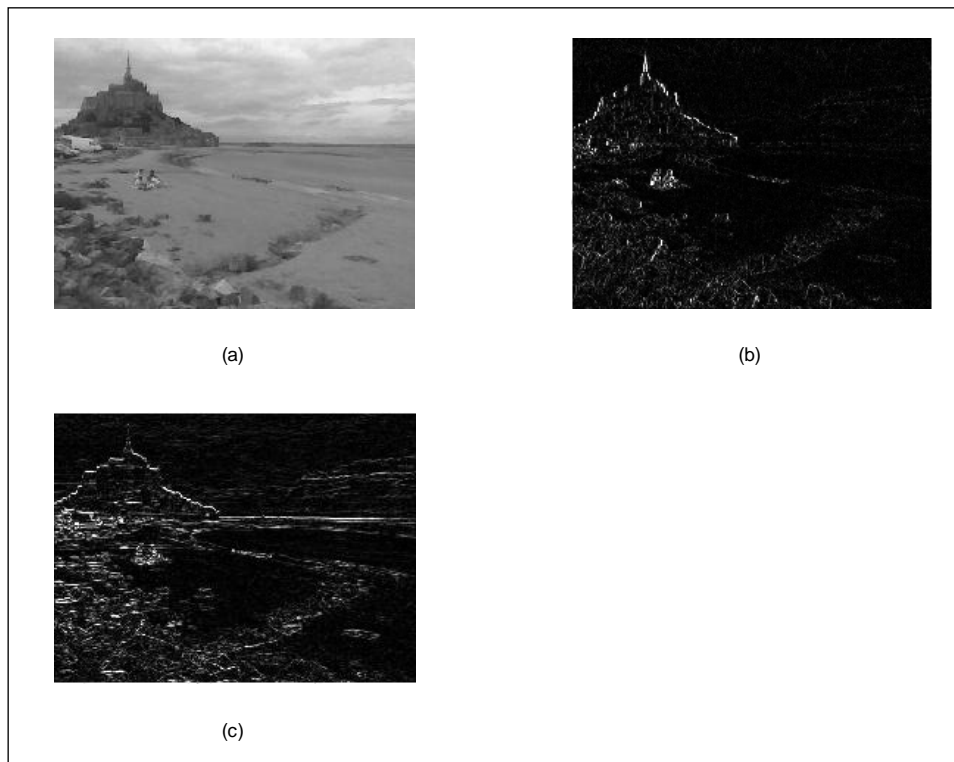


Figura 3-12, Efecto de la aplicación del operador de Sobel; (a) imagen original, (b) tras aplicar el operador horizontal, (c) tras aplicar el operador vertical.

3.4.2. Eliminación del *background*/Substracción de fondo

La substracción de fondo es una técnica que plantea la siguiente situación: la imagen tiene un trasfondo, *background* o sustrato, con cierta textura homogénea en toda captura que se toma. Además, el objeto o interés de la imagen se encuentra por encima de este sustrato y es identificable a ojos vista. Con estas condiciones impuestas, obtener la imagen deseada es tan sencillo como comparar la imagen deseada con la del sustrato y discriminar aquellos píxeles que mantengan un valor similar entre ambas imágenes.

El punto de este método es su sencillez, pues es comparable al concepto de “buscar las diferencias”. Pese a todo, presenta inconvenientes además de dificultades que se observan una vez implementado. No elimina las sombras producidas por el escáner (forman parte de la región diferente), y es muy susceptible a cambios de luz; en muchas ocasiones, la textura del sustrato no es homogénea o fácil de definir; y en menor medida puede identifica objetos inexistentes como sombras o elementos.



Figura 3-13, Escaneo de billete tras sustraer el trasfondo.

En el caso de que se desconozca la información del trasfondo, una alternativa es la aplicación de la detección de contornos. Cuando uno trabaja con un trasfondo se espera que no tenga cambios bruscos en su intensidad o nivel, que su textura sea “suave”.

Un ejemplo claro de aplicación de este método son las cámaras de video-vigilancia con detección de intrusos: estos aparatos, además de reproducir una imagen que puede o no ser grabada, disponen de un bucle retroalimentado de las imágenes tomadas que les sirve como referencia para la detección de cuerpos extraños.

4. Fundamentos matemáticos: PCA e ICA

El Análisis de Componentes Independientes (ICA) y Análisis de Componentes Principales (PCA)¹⁵ son dos métodos de Separación Ciega de Señales (abreviado a BSS¹⁶ en adelante). Estos métodos se caracterizan por su aplicación sobre grandes conjuntos de datos en los que se desconoce, u omite, todo conocimiento sobre la naturaleza de la señal o sus fuentes.

Concretamente, ICA y PCA son dos procedimientos que aportan cierta información sobre un conjunto de señales que se supone tienen cierta relación en cuanto a procesos físicos implicados. La información que aportan depende del método aplicado (PCA o ICA), pero generalmente diremos que extrae, de dicho conjunto de señales, un nuevo conjunto con cierta característica de independencia entre ellas. Este nuevo conjunto de señales se consigue bajo la condición que pueden combinarse para formar cualquier señal del conjunto original del que se extrajeron, sin importar que clase de señales fueran o sean tras el proceso.

En este documento se presenta ICA como un método para separar “mezclas” o mixturas de señales en sus señales originales y, en los capítulos siguientes, PCA como un método para reconocer patrones sobre imágenes digitales.

4.1. Introducción al Análisis de Componentes Independientes (ICA)

“Dos procesos físicos no relacionados son estadísticamente independientes.”

Este enunciado da pie al planteamiento en el que se fundamenta ICA. El objetivo de ICA (al igual que de los métodos BBS en general) es obtener un conjunto de señales independientes entre sí, y es con esa intención que formula un enunciado que, aunque parece acertado, resulta impreciso: “Si de una señal que es una mixtura (mezcla) de señales se extrae un conjunto de señales que son estadísticamente independientes, estas corresponden a diferentes procesos físicos”.

4.1.1. Ejemplo: La premisa de la fiesta

Enunciado por Christian Jutten y Jeanny Héroult en 1985¹⁷, se utiliza este caso como ejemplo para facilitar el entendimiento de la metodología y su utilidad.

Supongamos dos micrófonos separados que graban dos oradores de forma simultánea. Cada uno de estos micrófonos grabará una señal de audio distinta (que será una composición de ambas oratorias) influenciada por la distancia a la que se encuentran y otros factores. ¿Cómo obtener, de estos micrófonos, las señales de audio que corresponden a cada orador?

Desde el punto de vista de ICA presentado anteriormente; dos voces distintas corresponden a dos procesos físicos distintos, con lo que ICA supone que estas pueden ser reconstruidas a partir de las señales-componente que se obtienen al aplicar el método.

¹⁵ *Independent Component Analysis & Principal Component Analysis.*

¹⁶ *Blind Signal Separation.*

¹⁷ Wikipedia: es.wikipedia.org/wiki/An%C3%A1lisis_de_componentes_independientes.

4.2. Separación Ciega de Señales (BBS)

En esta sección se presentarán las estrategias que constituyen los métodos mencionados (ICA y PCA), además de la estrategia generalizada para los métodos de BSS. Se evitará profundizar demasiado en el aspecto técnico evitando mencionar aspectos y conocimientos no necesarios para comprender las variables clave para la implementación de los métodos.

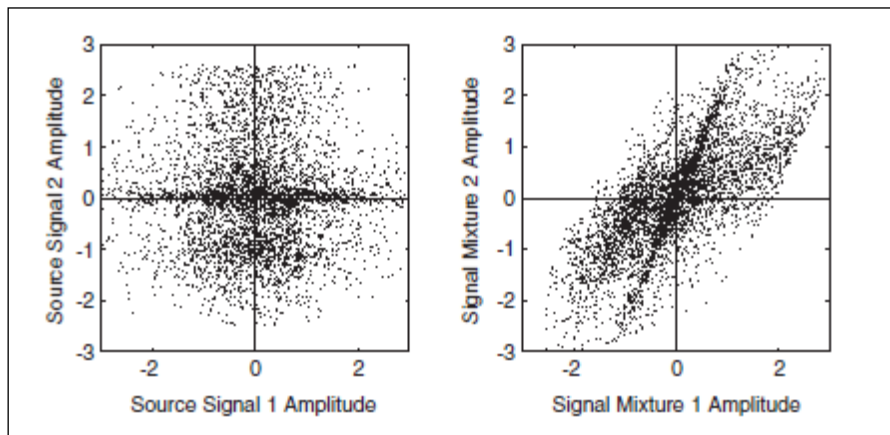


Figura 4-1, A la izquierda, distribución de dos variables aleatorias independientes. A la derecha, distribución de dos variables con dependencia.¹⁸

4.2.1. Señales fuente y mixturas de señales

Distinguiremos, en todo momento, dos clases de señal:

- 1- *Source signal*, o señal fuente; corresponden a las señales que se intentan extraer y forman las mixturas de señales.
- 2- *Signal mixture*, o mixtura de señal; corresponden a la suma, de varias señales fuente y son el punto de partida en el problema que se plantea.

Es importante distinguir entre ambas clases de señal no solo porque son el qué y porqué de la cuestión, también son señales con claras características que las diferencian. Y son estas diferencias un punto clave en los métodos BSS, ya que hallar el resultado depende mayormente de ellas. Distinguimos tres características o aspectos en los que se diferencian las *source signal* de las *signal mixture*:

- Independencia: Como se ha mencionado anteriormente, las *source signal* disponen de independencia estadística entre ellas. Se sobreentiende que las *signal mixture* a las que se aplica BBS no son independientes (Figura 4-1).
- Normalidad: Si se observa el histograma de una *source signal*, este tiene forma “puntiaguda”. Por el contrario, una *signal mixture* suele tener una forma gaussiana o normal (Figura 4-2).
- Complejidad: La complejidad temporal de una *signal mixture* será mayor, o al menos igual, a la de su *source signal* más simple o sencilla.

¹⁸ Figura extraída de Stone, James V., *Independent component analysis : a tutorial introduction*, ISBN 0-262-69315-1.

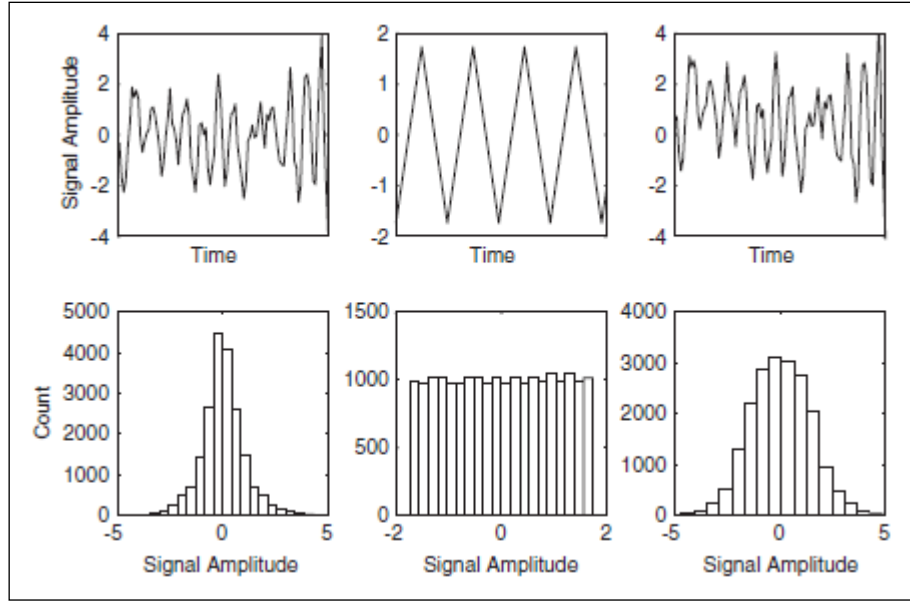


Figura 4-2, Señal y forma; Izquierda, señal super-gaussiana; Centro, señal dientes de sierra; Derecha, señal gaussiana.¹⁹

4.2.2. Estrategia general para BSS

Como premisa general, todos los métodos BBS buscan el mismo fin: obtener un conjunto de señales que responden a ciertas características basándose en un segundo conjunto de señales que se supone son mezclas del primer grupo.

Se nombra entonces s al un conjunto de señales originales que son independientes entre sí. Se nombra x al conjunto de señales que son mezclas de las señales s . La Ecuación 4.1 y la Ecuación 4.2 muestran un ejemplo de su notación en forma matricial. El valor simbólico k representa el número de señales disponibles y recuperables; N es el número de muestras de cada señal.

$$s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_k \end{pmatrix} = \begin{pmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,N} \\ s_{2,1} & s_{2,2} & & s_{2,N} \\ \vdots & & \ddots & \vdots \\ s_{k,1} & s_{k,2} & \cdots & s_{k,N} \end{pmatrix} \quad \text{Ecuación 4.1}$$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,N} \\ x_{2,1} & x_{2,2} & & x_{2,N} \\ \vdots & & \ddots & \vdots \\ x_{k,1} & x_{k,2} & \cdots & x_{k,N} \end{pmatrix} \quad \text{Ecuación 4.2}$$

Se interpreta la Ecuación 4.1 y la Ecuación 4.2 de la siguiente forma: la variable s o x , es un conjunto de vectores s_i y x_i , donde i representa cada uno de los k vectores que conforman la variable. Estos vectores, alineados en fila, son señales discretizadas de hasta N muestras ordenadas. Así pues, $x_{i,n}$ es la n -ésima muestra de la i -ésima señal de x .

Para una composición lineal, los conjuntos de señales s y x se relacionan según se muestra en la Ecuación 4.3 y Ecuación 4.4. A y W son dos matrices que multiplican a las señales s y x

¹⁹ Figura extraída de Stone, James V., *Independent component analysis : a tutorial introduction*, ISBN 0-262-69315-1.

respectivamente. Se entiende que tanto A como W son matrices cuadradas de orden k . La matriz A se corresponde a la matriz de mezcla o de mixtura, que representa el modo en que las señales fuente se mezclan para dar lugar a las señales mezcla. La matriz W se corresponde a la matriz recuperadora, que obtiene las señales originales a partir de las mezclas.

$$x = A \cdot s \quad \text{Ecuación 4.3}$$

$$s = W \cdot x \quad \text{Ecuación 4.4}$$

Entre A y W existe relación. Cuando se substituye la Ecuación 4.3 en la Ecuación 4.4 (o a la inversa), se obtiene la Ecuación 4.5 y la Ecuación 4.6. I_k es la matriz identidad de orden k , y señala que A y W son matrices inversas (Ecuación 4.7).

$$s = W \cdot (A \cdot s) = W \cdot A \cdot s \quad \text{Ecuación 4.5}$$

$$I_k = W \cdot A \quad \text{Ecuación 4.6}$$

$$A = W^{-1} \leftrightarrow W = A^{-1} \quad \text{Ecuación 4.7}$$

4.2.3. Matriz de mezcla y matriz recuperadora

Para facilitar el entendimiento, supondremos el caso del enunciado ejemplo; es decir, se especifica que el número de señales implicadas es $k = 2$. Esto simplifica la Ecuación 4.1 y Ecuación 4.2:

$$s = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \begin{pmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,N} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,N} \end{pmatrix} \quad \text{Ecuación 4.8}$$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,N} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,N} \end{pmatrix} \quad \text{Ecuación 4.9}$$

Donde A y W son dos matrices cuadradas de orden $k = 2$, y $a_{f,c}$ y $w_{f,c}$ los elementos singulares que las forman (f para el número de fila y c para el número de columna).

Para componer una señal x_1 se supone que las señales s_1 y s_2 han sido mezcladas, de forma que la recepción de estas parece una única señal. Seguidamente, se tiene en cuenta que esta composición es poco probable que sea tan simple como la suma de ambas señales y se propone un factor que modifica la amplitud de las señales antes de sumarlas (Ecuación 4.10). A este factor se le nombra *mixing factor* o factor de mezcla.

$$x_1 = a \cdot s_1 + b \cdot s_2 \quad \text{Ecuación 4.10}$$

$$x_2 = c \cdot s_1 + d \cdot s_2 \quad \text{Ecuación 4.11}$$

Para seguir con el enunciado ejemplo sobre la composición de voces, cada uno de estos factores puede asociarse a la atenuación de la voz a causa de la distancia al micrófono o dispositivo de grabación.

Siguiendo la forma de la Ecuación 4.10 se llega a la conclusión de que la señal mezcla x_2 se obtiene de forma similar (como la suma de las mismas señales s_1 y s_2), pero con distintos factores de mezcla (Ecuación 4.11). Cuando esto se compara con la Ecuación 4.3 y Ecuación 4.4, se obtiene que los factores de mezcla a, b, c , y d de la Ecuación 4.10 y Ecuación 4.11 se corresponden con los elementos de la matriz de mezcla A .

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \cdot \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \quad \text{Ecuación 4.12}$$

Del mismo modo, se define un proceso similar para el proceso de recuperación: Para obtener cada una de las distintas señales fuente s_i generalmente debe obtenerse una composición lineal de las señales x_j . Digamos pues, que una señal fuente s_1 puede obtenerse mediante la suma de las señales x_1 y x_2 multiplicadas por un *unmixing factor* o factor de separación, de forma respectiva tal y como se muestra en la Ecuación 4.13. Al igual que antes, estos factores se corresponden con los elementos de la matriz recuperadora W (Ecuación 4.15).

$$s_1 = \alpha \cdot x_1 + \beta \cdot x_2 \quad \text{Ecuación 4.13}$$

$$s_2 = \gamma \cdot x_1 + \delta \cdot x_2 \quad \text{Ecuación 4.14}$$

$$\begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = W \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{Ecuación 4.15}$$

Llegado este punto la discusión es la siguiente: como obtener dichos factores recuperadores $w_{i,j}$ de las distintas señales x_1 y x_2 . A partir de aquí, cada método de BSS sigue una estrategia distinta. Por ejemplo, ICA supone una matriz recuperadora con la que obtiene unas señales y , de las que evalúa la entropía. Otro método, nombrado como *Projection Pursuit*, separa las filas de la matriz recuperadora en vectores y busca que estos sean ortogonales a unos ejes del dominio de las señales x . En cualquier caso, estos métodos son procesos iterativos, que suponen una posible solución y evalúan por sus características si la solución supuesta es la “mejor posible”.

$$x = A \cdot s = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \cdot \begin{pmatrix} s_{1,1} & s_{1,2} & \cdots & s_{1,N} \\ s_{2,1} & s_{2,2} & \cdots & s_{2,N} \end{pmatrix} \quad \text{Ecuación 4.16}$$

$$s = W \cdot x = \begin{pmatrix} \omega_{1,1} & \omega_{1,2} \\ \omega_{2,1} & \omega_{2,2} \end{pmatrix} \cdot \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,N} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,N} \end{pmatrix} \quad \text{Ecuación 4.17}$$

4.3. Método ICA

“Dado un conjunto de señales y tal que $y = W \cdot x$, donde W es la matriz recuperadora, y una función invertible multivariable g de y que haga $Y = g(y)$; se buscará una W que maximice la entropía de Y .”

El método de análisis de componentes independientes se sostiene en la búsqueda de la independencia estadística de las señales obtenidas y su forma de señal no-gaussiana. Se recuerda la primera de las tres propiedades que caracterizan y diferencian a una *source signal* de una *signal mixture*: “las *source signal* disponen de independencia estadística entre ellas”.

Hasta el momento, se ha mencionado que la independencia estadística implicaba que una señal dentro de un conjunto no podía formarse como combinación de cualquier otra que ella misma. Para definir de forma numérica la independencia entre las señales, ICA utiliza las funciones densidad de probabilidad (abreviado a *pdf*²⁰ de ahora en adelante) del conjunto de señales y de las señales que lo forman.

4.3.1. Función Densidad de Probabilidad (*pdf*)

La *pdf* de una variable describe la probabilidad relativa de que dicha variable tome cierto valor. La *pdf* de una variable es una función positiva y está definida entre el valor mínimo y máximo que toma la variable x . Su expresión no es general y depende directamente de su forma.

Se define junto a la *pdf* otra función sobre la probabilidad de una variable. La llamada Función Densidad Acumulada (*cdf*²¹, en sus siglas en inglés) se corresponde como una expresión de alternativa de la *pdf* que puede resumirse en la probabilidad de que la variable tome cierto valor o uno menor. Para una variable discreta, esto se calcula mediante la suma de las probabilidades individuales de los valores de x menores a x_0 . La expresión continua se expresa según la integral de la probabilidad individual, y como la expresión derivada de *pdf*.

$$cdf(x) = p(x \leq x_0) = \sum_{x_0}^x p(x) \quad \text{Ecuación 4.18}$$

$$cdf(x) = \int_{x_0}^x p(x) dx \quad \text{Ecuación 4.19}$$

$$pdf(x) = cdf'(x) \quad \text{Ecuación 4.20}$$

²⁰ Probability Density Function.

²¹ Cumulative Density Function.

La *cdf* es por definición, una función monótona creciente con valor inicial 0 y valor final unitario.

El papel de estas funciones en ICA es como identificador de la forma de las señales independientes que se buscan. Sabemos que la *pdf* de una señal fuente tiene una forma concreta, y se deduce que su *cdf* también tiene forma concreta. Para la aplicación del método, es más interesante la *cdf* por su monotonía creciente, que asigna un valor singular entre 0 y 1 a cada valor de la variable.

En adelante se verá como la forma que toman las funciones *pdf* y *cdf* nos permite calcular la entropía de una señal y, al maximizarla, obtener la matriz recuperadora W que recupera el conjunto de señales s que son señales fuente.

4.3.2. Entropía con una variable

Se ha indicado que el método de ICA se sustenta en la independencia de las señales, y que se busca una forma de cuantificar esta propiedad. En realidad, la independencia no es cuantificable, por ello ICA utiliza una medida de la incertidumbre de la señal, llamada entropía.

“La entropía es una medida de la uniformidad de la distribución de un conjunto delimitado de valores”²². Si se entiende que la entropía representa la incertidumbre, la uniformidad de la distribución suena a un concepto completamente opuesto. Cuando se dispone del diagrama de dispersión de dos señales independientes, sus puntos están dispersos (valga la redundancia) en el plano. A medida que esta dispersión por el plano es más uniforme, el valor de entropía aumenta. Es al maximizar esta entropía que el método ICA halla los elementos de la matriz recuperada, y que recuperan un conjunto de señales independientes entre sí.

La entropía se responde a la Ecuación 4.21 y se simboliza como $H(z)$, donde z es el suceso o variable en cuestión.

$$H(z) = - \sum_i^n p_z(z_i) \log(p_z(z_i)) \quad \text{Ecuación 4.21}$$

$$H(z) = - \int_z p_z(\zeta) \cdot \ln(p_z(\zeta)) d\zeta \quad \text{Ecuación 4.22}$$

Cuando se trata de eventos con una probabilidad uniforme (por ejemplo una señal discreta de N muestras) p_z es contante y se expresa como N^{-1} . Entonces la Ecuación 4.21 queda se escribe como la Ecuación 4.23.

$$H(z) = - \frac{1}{N} \sum_t^N \log(p_z(Z^t)) \quad \text{Ecuación 4.23}$$

²² Stone, capítulo 7: Independent Component Analysis.

Mencionar que la Ecuación 4.21 es máxima cuando el valor p_z es constante para los n eventos posibles. Esto se traduce a que la entropía es máxima cuando una variable aleatoria tiene una distribución uniforme.

El concepto de entropía se aplica en la recuperación de señales de la siguiente forma: Se supone una función g que es monótona creciente, y que se aplica sobre un conjunto de señales $y = W \cdot x$. Al hacerlo se obtendrá $Y = g(y)$. La función g se define de forma que Y tenga una distribución uniforme y máxima entropía cuando la señal recuperada y tenga una pdf que se corresponda a la pdf de una señal fuente.

Tras la discusión del enunciado anterior, se puede obtener que el valor de entropía de la señal Y depende de la pdf de y tal y como se muestra en la Ecuación 4.24. Donde p_y es la pdf de la señal y y p_s la pdf de la señal fuente. El punto conflictivo de este apartado es la función g ; que es una función definida para el método. Esto quiere decir que podría utilizarse la función g para recuperar una señal distinta a una señal fuente. Para el caso en cuestión, g' se corresponde a p_s , por lo que, g es la función cdf de una señal fuente.

$$H(Y) = -\frac{1}{N} \sum_t \log \left(\frac{p_y(y^t)}{p_s(y^t)} \right) \quad \text{Ecuación 4.24}$$

$$g = \int g'(y) dy = \int pdf(y) dy = cdf(y) \quad \text{Ecuación 4.25}$$

Entonces, si se obtiene que Y tiene dicha máxima entropía, entonces el conjunto de señales y tendrá independencia estadística, y se corresponderá al conjunto s .

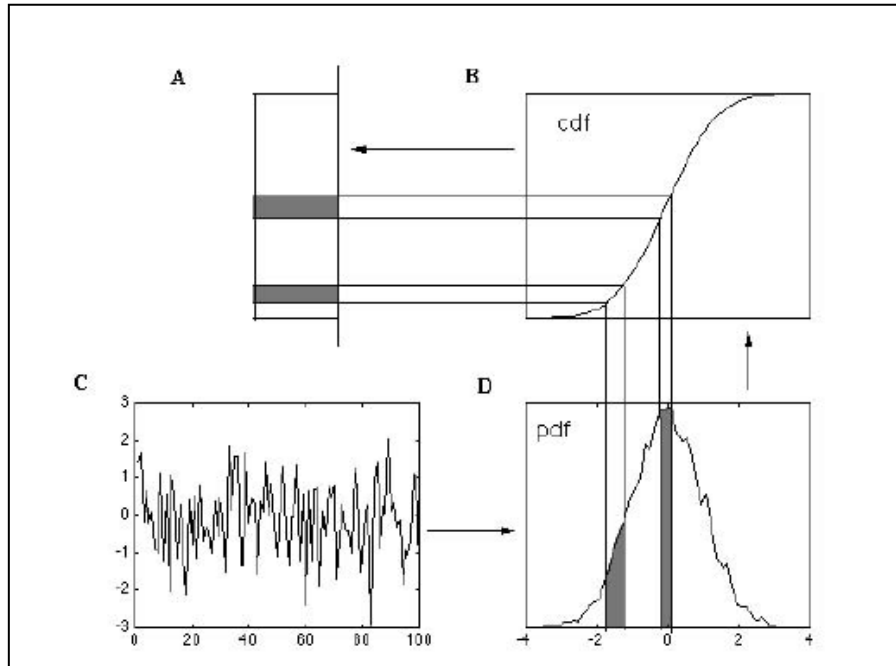


Figura 4-3, Esquema sobre la interpretación de la función g .

4.3.3. Independencia del conjunto y distribución marginal

Cuando se trata con un conjunto de señales, su función densidad de probabilidad (*pdf*) debe mostrar tanto la influencia de la *pdf* individual de cada señal y la *pdf* de la dependencia entre estas. Por suerte, se está tratando con señales que se desea y suponen independientes entre ellas. Por ello, la función distribución conjunta podrá expresarse como el producto de cada *pdf* individual (Ecuación 4.26). A estas funciones *pdf* de señales que forman el conjunto se les llama “Distribución marginal de variable”. La Ecuación 4.27 muestra el caso de $k = 2$.

$$p_s(s) = \prod_{i=1}^k p_{si}(s_i) \quad \text{Ecuación 4.26}$$

$$p_s(s_1, s_2) = p_{s1}(s_1) \cdot p_{s2}(s_2) \quad \text{Ecuación 4.27}$$

4.3.4. Entropía con múltiples variables

Siguiendo el ejemplo anterior, la entropía de un conjunto de variables se conoce como entropía conjunta y puede entenderse como la distribución de los puntos que forman el conjunto cuando estos se dibujan unos contra otros.

$$H(Z_1, Z_2) = \int_{\zeta_1} \int_{\zeta_2} p_z(\zeta_1, \zeta_2) \cdot \log(p_z(\zeta_1, \zeta_2)) d\zeta_1 d\zeta_2 \quad \text{Ecuación 4.28}$$

$$H(\mathbf{Z}) = \int_{\mathbf{Z}} p_z(\mathbf{z}) \cdot \log(p_z(\mathbf{z})) d\mathbf{z} \quad \text{Ecuación 4.29}$$

La expresión de la entropía para un conjunto de dos variables se muestra en la Ecuación 4.29. Para abreviar próximas expresiones y mostrar que se considerará que el conjunto de señales (Z_1, Z_2, \dots, Z_N) se corresponde a \mathbf{Z} . Donde $p_z(\mathbf{Z}) = p_z(\zeta_1, \zeta_2)$ es la función distribución conjunta.

$$H(\mathbf{Z}) = - \sum_i^n p_z(\mathbf{Z}) \log(p_z(\mathbf{Z})) \quad \text{Ecuación 4.30}$$

Para un finito número de muestras del conjunto \mathbf{Z} la expresión de la entropía conjunta se corresponde a la Ecuación 4.31.

$$H(\mathbf{Z}) = - \frac{1}{N} \sum_i^N \log(p_z(\mathbf{Z}^t)) \quad \text{Ecuación 4.31}$$

Se reconsidera el planteamiento base del método aplicado a múltiples variables. Se considera la entropía del conjunto de señales $\mathbf{Y} = g(\mathbf{y})$, donde $\mathbf{y} = \mathbf{W}\mathbf{x}$ es el conjunto de señales extraído por la matriz recuperadora \mathbf{W} . Para un finito número de muestras del conjunto, se

expresará como la Ecuación 4.32. Tal como con la versión de una variable, la entropía del conjunto será máxima cuando la distribución del conjunto Y sea uniforme.

$$H(Y) = -\frac{1}{N} \sum_i^N \log(p_Y(Y^t)) \quad \text{Ecuación 4.32}$$

Siguiendo la misma línea de razonamiento que en apartado anterior, se obtiene que el valor de la entropía del conjunto Y depende de la distribución de y y de la distribución del conjunto de señales fuente que se esperan recuperar. Como en el caso de una variable, cuando la distribución de ambos conjuntos sea la misma, la entropía de Y será máxima.

$$H(Y) = -\frac{1}{N} \sum_t^N \log\left(\frac{p_y(y^t)}{p_s(y^t)}\right) \quad \text{Ecuación 4.33}$$

El siguiente paso es hallar una expresión para $p_y(y^t)$ que permita evaluar la Ecuación 4.33. Para ello, se define la probabilidad de p_y como una expresión de p_x . Se halla que la relación existente entre los conjuntos y y x es el valor absoluto del determinante de la matriz recuperadora W .

$$p_y(y) = \frac{p_x(x)}{\left|\frac{\partial y}{\partial x}\right|}; \quad \left|\frac{\partial y}{\partial x}\right| = |W| \quad \text{Ecuación 4.34}$$

Cuando se sustituye el resultado de la Ecuación 4.34 en Ecuación 4.33 se obtiene que la entropía de Y depende de la entropía de la distribución s , de la distribución x y del valor del determinante de la matriz W . Para poder simplificar la expresión, se expresa en forma de suma de logaritmos (Ecuación 4.36).

$$H(Y) = -\frac{1}{N} \sum_t^N \log\left(\frac{p_x(x^t)}{|W| \cdot p_s(y^t)}\right) \quad \text{Ecuación 4.35}$$

$$H(Y) = \frac{1}{N} \sum_t^N \log(p_s(y^t)) + \log(|W|) - \frac{1}{N} \sum_t^N \log(p_x(x^t)) \quad \text{Ecuación 4.36}$$

Según la lógica del problema, la distribución conjunta de x es una constante, por ello se define una expresión simplificada de la entropía de Y , $h(Y)$ que no tiene en cuenta la distribución conjunta de x (Ecuación 4.37).

$$h(Y) = \frac{1}{N} \sum_t^N \log(p_s(y^t)) + \log(|W|) \quad \text{Ecuación 4.37}$$

Llegado a la expresión de la Ecuación 4.37, solo queda desglosar la distribución conjunta de \mathbf{y} . Si como se espera y desea, el conjunto de señales \mathbf{y} es independiente, entonces su distribución conjunta puede expresarse como el producto de sus distribuciones marginales, que a su vez, puede expresarse como la suma de N logaritmos, tal y como muestran la Ecuación 4.38 y la Ecuación 4.39.

$$h(\mathbf{Y}) = \frac{1}{N} \sum_t \log \left(\prod_{i=1}^k p_{si}(y_i) \right) + \log(|W|) \quad \text{Ecuación 4.38}$$

$$h(\mathbf{Y}) = \frac{1}{N} \sum_i^k \sum_t^N \log(p_s(y_i)) + \log(|W|) \quad \text{Ecuación 4.39}$$

Con la expresión de la Ecuación 4.39 solo resta expresar la distribución marginal $p_s(y_i)$ en función de la forma de los vectores w_i que forman la matriz recuperadora W .

Llegados este punto, el inciso sobre el método ICA no se extenderá más. Los puntos restantes son la implementación del método de búsqueda de los elementos de W mediante un gradiente adecuado. En esta sección, se buscaba discutir y mostrar los elementos clave implicados en el cálculo de las ICA como es la maximización entropía de la señal \mathbf{Y} .

4.4. Análisis de Componentes Principales (PCA)

Históricamente, PCA es un método que se ha utilizado extensamente con los problemas que ahora se plantean con ICA. La principal diferencia entre ambos es que, si ICA obtiene un conjunto de señales con *pdf* no-gaussiana e independientes, PCA busca sencillamente un conjunto de señales con *pdf* gaussianas no-correlacionadas.

Aunque PCA no busca la independencia de las señales que obtiene, existe el caso en que estas si los son; cuando las señales estadísticamente independientes son, además, gaussianas.

Si se comparan los momentos, ICA tiene un mayor nivel de independencia respecto a PCA. Los conjuntos de señales obtenidos por PCA deben cumplir la Ecuación 4.40, que relaciona la expectación del conjunto con la de las señales individuales. ICA por el contrario, debe cumplir la Ecuación 4.41 para cualquier, momento de q y p .

$$E[y_1, y_2] = E[y_1] E[y_2] \quad \text{Ecuación 4.40}$$

$$E[y_1^p, y_2^q] = E[y_1^p] E[y_2^q] \quad \text{Ecuación 4.41}$$

PCA además de hallar una matriz recuperadora de señales que son no-correlacionadas, organiza las señales extraídas según su variabilidad, o su aportación en potencia/energía a las mixturas de las que se han extraído. Así aquellas señales con mayor varianza asociada se consideran más importantes en el momento de mezclar las señales.

Concretamente, PCA obtiene un conjunto ordenado de señales gaussianas tal que cada señal aporta una menor proporción a la variabilidad de las mixturas. La naturaleza de las señales no-correlacionadas asegura que las aportaciones de distintas señales no se solapan ni anulan magnitudes de variabilidad en el conjunto de señales-mixturas. A esta variabilidad se le conoce formalmente como varianza.

4.4.1. Estrategia de PCA

Al igual que con otros métodos, PCA comienza proponiendo la separación de la matriz recuperadora en vectores que recuperan una única señal, conocida como Componente Principal o PC. Estos vectores, expresados como w_i se supondrán de modulo unitario, y el valor de varianza asociado a su PC se simboliza con λ^{23} . A diferencia de los vectores w_i propuestos por ICA, los vectores de PCA son mutuamente perpendiculares. Esta característica es la que asegura la no-correlación entre las componentes principales.

$$y = \lambda W_{PCA} \cdot x \quad \text{Ecuación 4.42}$$

Así pues, PCA buscará un vector w_1 rotándolo en un espacio de M dimensiones (cambiando su orientación) y que dado el producto de este con el conjunto de señales fuente, se obtenga una variable y con varianza σ_1^2 máxima y que no puede ser excedida por ningún otro vector w_i . La búsqueda de este vector w_1 se realiza rotando el nuevo vector en el mismo espacio de M dimensiones. El resto de vectores w_i serán, además de unitarios, ortogonales a w_i y al resto de vectores ya calculados.

$$y_1 = \lambda_1 w_1 \cdot x \quad \text{Ecuación 4.43}$$

$$|w| = \sqrt{w_1^2 + w_2^2 + \dots + w_n^2} = 1 \quad \text{Ecuación 4.44}$$

Un apunte sobre la última de las componentes es que esta no completa la varianza; la suma de varianzas de las componentes no tiene que ser igual a la de las señales de las que se extrajeron, y generalmente no lo hacen. La última componente es sencillamente aquella con menor peso o importancia dentro del grupo.

4.4.2. PCA y señales fuente

El método PCA no obtiene un conjunto de señales fuente. Cada una de sus componentes no se guía por la independencia sino por la no-correlación entre ellas y por su máxima varianza. Se espera que aparte de la primera componente principal, el resto no se corresponda con ningún proceso físico. El resto de componentes entre la primera (aquella con mayor valor de varianza) y la última (aquella con menor varianza), se definen literalmente como “componentes con varianza menor a la de la componente principal”.

²³ λ se especifica como un escalar asociado al vector w correspondiente. En la Ecuación 4.42, λ es una matriz diagonal, donde cada elemento no-cero es el valor λ_i correspondiente al vector w_i .

4.4.3. PCA como filtro

Supongamos un gran conjunto de datos al cual aplicar cualquier transformación o análisis. PCA propone utilizar su método para reducir la envergadura de estos datos filtrando aquellos con menor importancia. Es decir, aplicar PCA y reconstruir el conjunto inicial ignorando las componentes de menor importancia.

Esta aplicación no es aliena a dichas funciones, incluso en casos de aplicación de ICA se aplica como filtrado de datos innecesarios o de poca relevancia. En análisis de señales con ruido, tiene gran efectividad, ya que las señales de ruido suelen por definición, tener poca aportación energética a la señal.

4.5. Uso de PCA como método de reconocimiento

Recordemos que la diferencia principal entre ICA y PCA es el grado de independencia del resultado: ICA aportaba señales que eran estadísticamente independientes entre sí, y que podíamos asociarlas a procesos físicos diferentes; PCA por el contrario se limita a señales no correlacionadas entre sí y que no podían asociarse a procesos físicos concretos o distintos.

¿Por qué entonces escoger PCA cuando ICA muestra una mayor ventaja en cuanto a independencia? Principalmente por su manejabilidad; para el estudio que se realiza no buscamos elementos con gran rango de independencia, sino elementos “esenciales/importantes” y aquellos con menor nivel de esta característica para poder discriminarlos. Las señales obtenidas de PCA aportan estas características; estas se hallan y organizan según su aportación a la varianza del conjunto de señales-mixtura. Para el problema que planteamos, esto implica que si discriminamos el último de estos elementos ordenados, que se corresponde según el orden con la última de las señales extraídas, se puede volver a obtener la mixtura de las que se extrajeron con error mínimo.

Si interpretamos el caso aplicado, PCA descompondrá nuestro conjunto de imágenes en nuevas imágenes que se interpretan como capas de cada una de ellas. Debido al método, sabremos que cada una de estas capas está presente en cada una de las imágenes, y que además algunas de ellas tienen mayor importancia que otras. Aquellas que tengan menor varianza dentro del conjunto serán descartadas, pues pueden asociarse a señales de ruido o al efecto de sombras y detalles menores.

También se debe aplicar este filtrado a la imagen que ha de ser reconocida. Debido al efecto de filtrado de ambos, el conjunto de imágenes y de la imagen a reconocer, se espera que el proceso de reconocimiento sea más claro y por ello dé resultados más fiables.

4.6. Resumen

Durante el capítulo se han presentado los métodos BSS, Separación Ciega de Señales, de los cuales ICA y PCA son casos particulares. Se han presentado la estructuración de estos y condiciones particulares que los difieren; el planteamiento y fundamento en que se basan desarrollo de cada método particular hasta alcanzar una comprensión extensa pero no exhaustiva de este.

Como conclusión, se ha discutido la elección del método que se utilizará para los capítulos posteriores. Haciendo hincapié en la interpretación de los resultados intermedios obtenidos en el proceso (las componentes principales), y en los posibles efectos directos.

5. Implementación de software

5.1. Revisión del grupo de datos

El grupo de datos que se ha provisto para el proyecto consta de una numerosa cantidad de imágenes de distintos billetes de Euro bajo cuatro espectros electromagnéticos, entre ellos se puede formar el espectro de la luz visible. Las imágenes de estos billetes están clasificadas según el valor y año de puesta en circulación para aquellos billetes de nuevo diseño. Para cada una de estas categorías existen cuatro diferenciaciones según la orientación del billete en el momento de ser escaneado; Anverso (A), Anverso invertido (AI), Reverso (R), y Reverso invertido (RI). Se dispone, para cada una de las $(7+4) \times 4$ clasificaciones, de un total de 50 muestras irradiadas con luz visible y captadas como imagen en escala de grises.

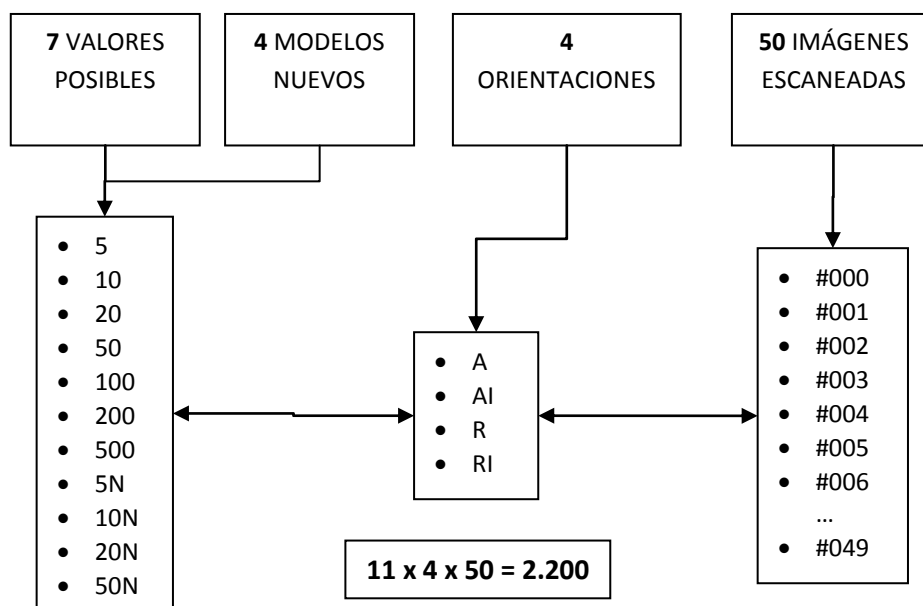


Figura 5-1, Esquema de la organización de las imágenes escaneadas.

5.1.1. Grupo de datos de entrada

Todas las muestras disponibles conforman el grupo de datos de entrada. Desde un principio, su uso tendrá dos fines:

- Construir un modelo de referencia que represente una idealización de las muestras. Se establece que para cada clase se construirá un modelo de referencia distinto.
- Utilizarlos de forma individual como una entrada de datos nuevos virtual con los que analizar resultados.

Las muestras se clasifican, como se ha dicho anteriormente, según once valores nominales (7 valores estándar más cuatro versiones actualizadas) y cuatro posiciones (Anverso, Reverso, etc.). La clasificación de las muestras de las que se dispone se refleja en su nombre de archivo. Un ejemplo posible es el nombre de archivo “50R#013” y “10NAI#032”. El primero se corresponde a la decimotercera muestra del billete de 50€ por el Reverso; El segundo se

corresponde a la trigésimosegunda muestra del billete de 10€ expedido en 2014²⁴ (N) por el Anverso Invertido.

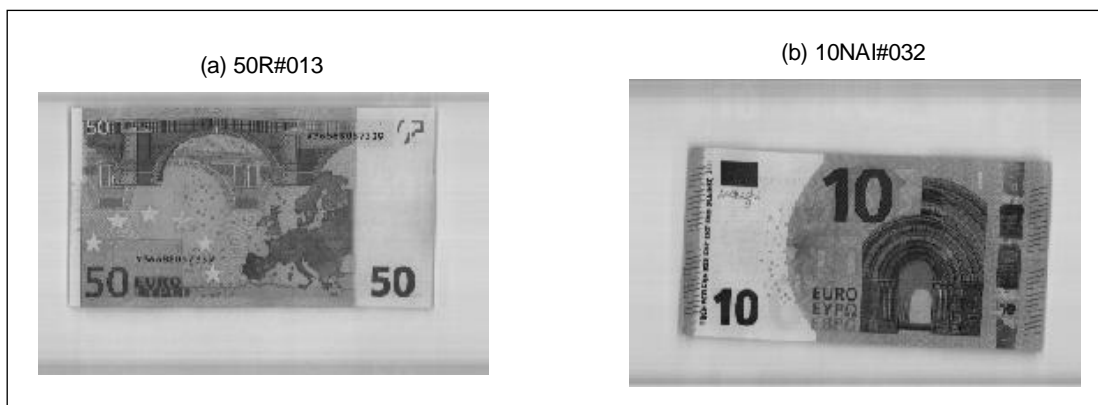


Figura 5-2, Imágenes originales de los ejemplos: (a) 50R#013, (b) 10NAI#032.

Esta clasificación y ordenación están presentes en los nombres de archivo de cada una de las muestras utilizadas, y se hace uso de estos nombres dentro de código para poder seleccionarlos y cargarlos al entorno de MATLAB. Con esto se quiere indicar que si bien los nombres de archivo forman parte del código, estos no se utilizan para identificar la muestra más allá de la fase de preparación del modelo de referencia y de la llamada de la entrada simulada de muestras en la fase de análisis.

La adquisición de la imagen, y su carga en el entorno figura y se resume en la página 2 del Apéndice P: Códigos de programa.

5.1.2. Datos de salida

Se diseñará el software para que este produzca como resultado de las operaciones un vector de valores entre 0 y 1. Se diseñará dicho vector para que existan tantos elementos como clases se hayan definido, y su valor deberá reflejar el parecido con las señales de referencia definidas para cada clase. Además de este vector de valores, como salida, el *software* aportará una identificación en forma de *string*²⁵ que indique, de forma explícita, a que clase el programa asigna la imagen entrada. Por último, y para cálculos externos, se comprobará la veracidad del método comparando la clasificación de la imagen con su identificación real. Si coinciden, se emitirá junto con el resto de variables de salida una variable con valor “True”. Por el contrario, si no coincidieran, un valor “False” será emitido en su lugar.

²⁴ Las versiones nuevas de los billetes de 5, 10, 20 y 50 euros fueron expedidas en el año 2013, 2014, 2015 y 2017 (año actual) respectivamente.

²⁵ Cadena de caracteres. Formato habitual de variable en informática.

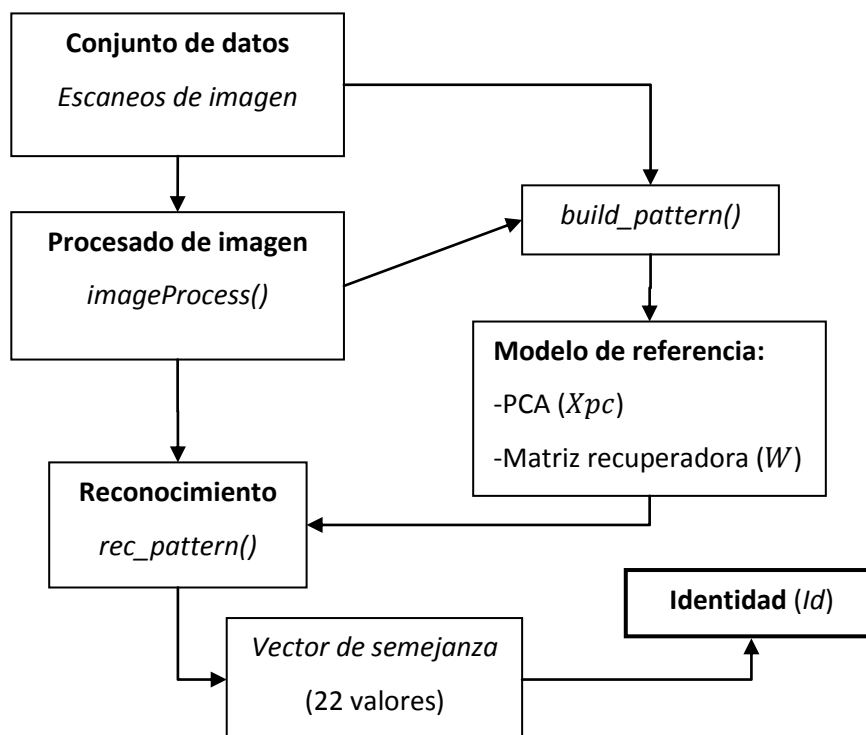


Figura 5-3, Diagrama de bloques que representa las etapas del *software*.

Las variaciones del código para la identificación de las muestras entradas, serán utilizadas generalmente para obtener valoraciones sobre la fiabilidad del método o su eficiencia.

5.1.3. Consideraciones sobre las muestras

La primera consideración serán las clases que se definirán para el *software*. Se definirán las clases basándose en la clasificación ya establecida en las muestras del grupo de datos. Se recuerda que el grupo de datos distingue 44 clases distintas en total (11x4); sin embargo, se diseñará el software para distinguir entre 22 clases únicamente. Esta consideración nace de la idea de unificar los grupos A y AI en uno, al igual que los grupos R y RI. La diferencia entre ambos grupos es un giro de 180°, que se incluirá en la lectura de la muestra tras su lectura desde el archivo.

En figuras e imágenes posteriores se intentará mostrar una visualización “cómoda” de las muestras. Es decir, que en algunos casos no se mostrará la imagen en la misma orientación a cuando se trata dentro del programa. Por ejemplo, todas las muestras se cargan en posición vertical como en la Figura 5-2. Sin embargo, la Figura 5-4, se muestra con una rotación anti horaria de 90°.

Para este proyecto se descarta, como parte esencial del procedimiento, la detección de falsificaciones. La existencia de la misma imagen bajo distintos espectros, y diferentes escaneos, permite que el proyecto se enfoque en la clasificación del valor nominal del billete. Se propone que si se haga uso de esas otras radiaciones si es de interés la identificación de falsificaciones.

5.2. Métodos implicados: Procesado de imagen

En esta sección se presentan las correcciones realizadas sobre la imagen y las decisiones que han llevado a ellas. Estos métodos han sido expuestos con anterioridad en el capítulo 3, Procesado de imagen digital, pero no se ha explicado su aplicación en el proyecto o sus resultados.

El objetivo de esta etapa del programa es la “mejora” de la imagen y localización del objeto de interés; a la que se referirá como la muestra, o imagen del billete. Así pues, primero se quiere corregir la inclinación de la muestra; segundo, se eliminará el fondo de la imagen, aislando la imagen del billete del resto. Para ello, se deberá localizar ciertos píxeles que sean referencia de la posición del objeto de interés, en concreto, se buscarán las esquinas del perímetro de la muestra.

Durante esta sección se mostrarán los cambios y procedimientos aplicados a una de las muestras para facilitar su visualización y entendimiento.



Figura 5-4, Imagen sin modificar de la muestra 100A#010.

El código asociado a estos procedimientos puede encontrarse en el Apéndice F: Códigos de figuras del capítulo 5. La muestra que se seguirá será la “100A#010”, y su imagen inicial se muestra en la Figura 5-4.

5.2.1. Eliminación de sombras y sustracción de fondo

Para corregir la inclinación y eliminar el fondo de la imagen es necesario localizar la región de la muestra que pertenece al objeto. Primeramente se aplica el método de sustracción de fondo para limpiar de la imagen los píxeles que no pertenecen al objeto. Para ello, se toma una

“traza” de la imagen sin influencias del objeto y con esta traza se recrea una simulación del fondo de la imagen (Figura 5-5). Tras la sustracción, se espera que todo pixel que no pertenece a la región del objeto tenga valor nulo.

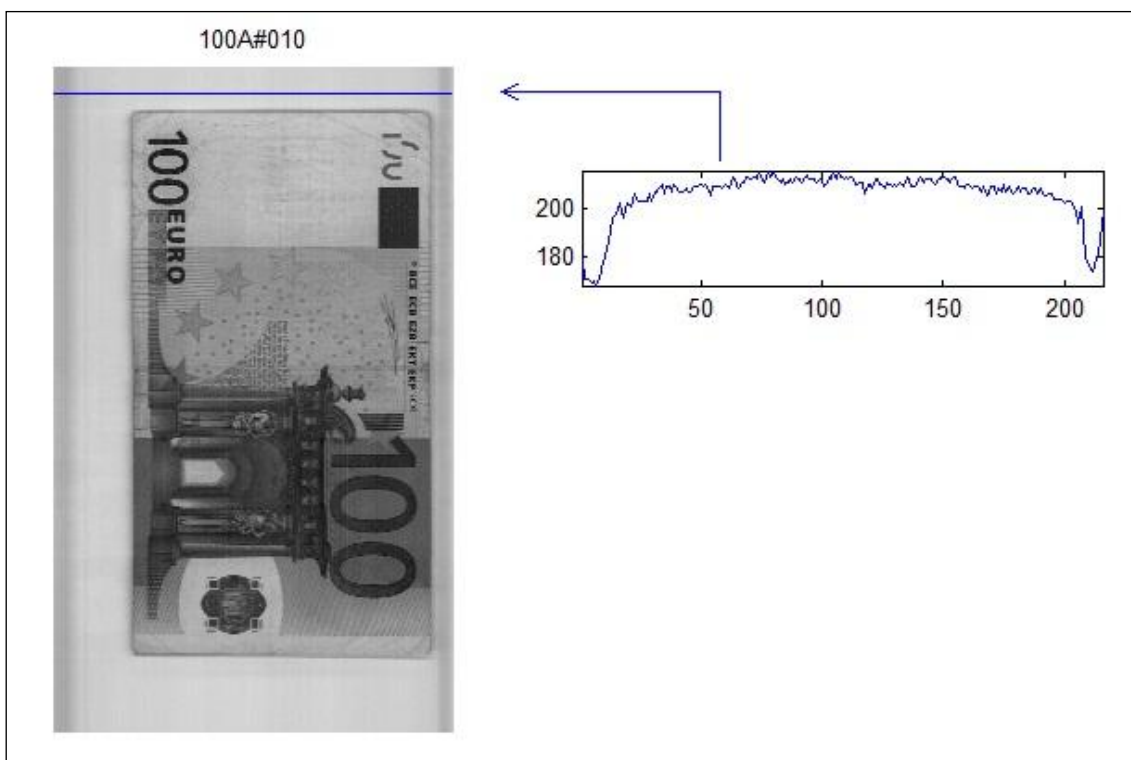


Figura 5-5, Selección de traza y su forma.

El proceso se resume en la página 2 del Apéndice P: Códigos de programa.

5.2.2. Detección de contorno

Para la detección del contorno, se propone utilizar un operador de Sobel sobre una imagen binaria donde se destaque la región de la imagen donde se localiza el objeto. Para adquirir la imagen binaria, se llevan a cabo varias modificaciones sobre el nivel de gris de la imagen que permitan diferenciar las regiones de la imagen según un valor umbral. Una vez aplicado el operador de Sobel a la imagen binaria, la imagen resultante es un resalte del contorno del objeto sobre un fondo negro.

Para localizar el contorno, se aprovecha el conocimiento previsto sobre su forma para localizar diversos de sus puntos. A partir de la imagen del contorno, se hallan las expresiones que definen las rectas que lo delimitan y, a su vez, estas marcan las cuatro esquinas del objeto.

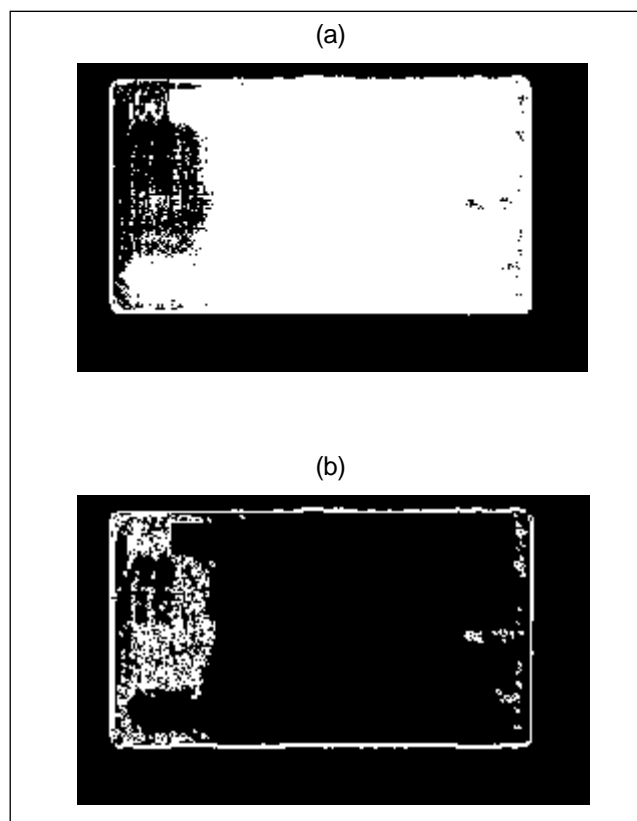


Figura 5-6, (a) imagen binaria donde se puede diferenciar el objeto y (b) contorno del objeto tras aplicar el operador de Sobel.

5.2.3. Corrección de la inclinación (Giro)

Una vez se ha localizado la región de la imagen que pertenece a la muestra, se desea corregir su orientación con el fin de aislar la muestra del resto de la imagen. En este punto, y debido principalmente a su geometría, la localización de las esquinas de la muestra permitirá el cálculo del ángulo a rotar. Debe tenerse en cuenta que tras la rotación, las dimensiones de la imagen cambian, y las coordenadas dejan de corresponderse entre ambas imágenes.

Una vez corregida la inclinación, la imagen puede recortarse a la región de la muestra, que será mayor o menor según la muestra en cuestión. Si la corrección de la inclinación ha sido exitosa, el proceso constará de recoger aquellos píxeles que se encuentren entre dos valores de fila y dos valores de columna. Estas columnas y filas marcan el perímetro de un objeto cuadrangular.

5.2.4. Reducción de la resolución espacial

Debido a las distintas dimensiones de las muestras una vez aisladas del resto de la imagen surge el siguiente problema: las muestras de menor tamaño tendrán señales más pequeñas y que diferirán demasiado de las más grandes. Esto se cree causará que al realizar el análisis de sus componentes suceda que las componentes con mayor importancia en ciertas clases no formen parte de las componentes del resto de clases.

En cualquier caso, para saltar este inconveniente se ajustaran las dimensiones de todas las clases de muestras a una misma. La solución permite dos opciones, aumentando las muestras de menor tamaño o disminuir las de mayor tamaño. Se decide reducir principalmente a que este curso de acción también aligera la carga de procesado.

Valor nominal	Lado mayor (L)	Desviación estándar del lado mayor	Lado menor (M)	Desviación estándar del lado menor	Proporción L/M
5	245,045	3,088	128,846	0,845	1,901
10	259,081	2,258	137,270	3,227	1,887
20	271,638	0,827	140,833	8,021	1,928
50	285,500	0,707	156,500	3,535	1,824
100	–	–	–	–	–
200	310,676	1,296	164,411	13,834	1,889
500	–	–	–	–	–

Figura 5-7, Tabla con las dimensiones medias de las imágenes en píxeles y su desviación estándar.

5.2.4.1. Efectos positivos y negativos

Aquí aparece el primer “selector”; la resolución de la muestra. Los selectores se definen en este proyecto como valores que tienen un impacto importante en los resultados del procedimiento. En esta ocasión el tamaño de la muestra influye en la carga de procesado, pero se cree que una menor resolución dificultará la distinción de la muestra entre distintas clases.

5.3. Modelos planteados

Para el proceso de clasificación de las muestras, estas se compararan con un conjunto de modelos existentes y que se consideran muestras representativas de las diferentes clases definidas para la clasificación. Siguiendo las indicaciones de la sección 5.1.3, se han definido 22 clases y sus respectivos 22 modelos.

En esta sección se presentan distintos métodos para construir un modelo de referencia para las clases que conforman el proceso de clasificación de muestras. El método de comparación pasa por al análisis de componentes y se explica en la sección que sigue.

5.3.1. Clasificación preliminar por dimensión

En esta sección se dice que se construirán modelos representativos de las clases especificadas, y que estos se compararán con las muestras a fin de comprobar si se corresponden. Aún así, nada impide que clasifiquemos las muestras por el tamaño de estas; se espera que dos muestras que pertenecen a la misma clase dispongan (una vez corregida la inclinación) de un tamaño similar.

En caso del euro, en que cada uno de los billetes es de tamaño distinto, esta clasificación es válida, pero inviable ya que no tiene en cuenta el contenido de la muestra. Para otros billetes esta idea es impensable ya que se mantiene el mismo tamaño para todas sus variaciones. Sin embargo, debido al procedimiento utilizado, se pierde esta referencia del tamaño, y su uso como clasificación preliminar no implica una gran carga para el programa.

El uso de este aspecto no está garantizado en el proyecto. Aún así se menciona por su sencilla implementación y uso.

5.3.2. Modelo idealizado

Para la construcción de este modelo se supondrá que para cada clase de billete existe un billete designado como “ideal”, del cual todos los demás son variaciones. Estas variaciones pueden ser ocasionadas por fatiga, el paso del tiempo, escritos sobre la muestra, rupturas, etc. Se defiende que las componentes extraídas de este modelo ideal no estarán influidas por las

variaciones mencionadas, con lo que se espera que sea el modelo más fiable en cuanto a reconstrucción completa.

El modelo se basa en la definición del patrón de un billete basándose en la intensidad de las regiones de este. La influencia de la luz puede ser un factor importante, con lo que distribuir los niveles de gris para alcanzar todos sus valores posibles es un procedimiento previo que gana importancia. Las posibilidades en cuanto a cómo construir pueden ser diversas, pero no se esperan grandes variaciones entre los diferentes métodos de construcción. Por su sencillez, se elige un proceso de acumulación y división. Es decir, se calcula la media aritmética de cada pixel de la muestra.

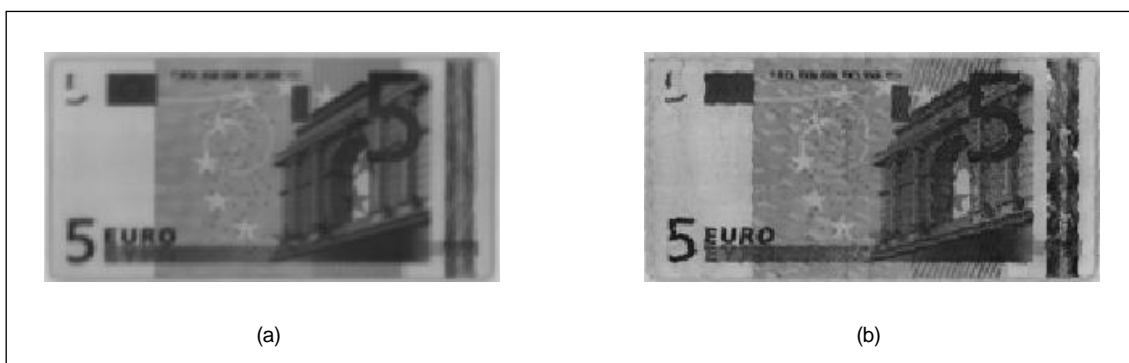


Figura 5-8, (a) Imagen idealizada (b) muestra sin modificar.

5.3.3. Modelo de Regiones De Interés (RDI)

La construcción del modelo seleccionando las regiones de interés es, tal vez, el método con mayor potencia para la clasificación de las imágenes del billete. El método en sí, supone ciertas complicaciones y requiere de un importante estudio de las muestras en general para la elección de las RDI. La selección de las regiones debe ser equivalente entre las diferentes clases para una clasificación exhaustiva, y deben escogerse para que caractericen tanto al billete en función del que se elige la RDI como lo diferencien del resto del grupo.

5.3.4. Modelo de Bordes

Entre la bibliografía del proyecto se menciona en diversas ocasiones a autores que implementaron métodos para el reconocimiento del papel basándose en sus bordes. Este método puede entenderse como una aplicación específica del modelo anterior 5.3.3, Modelo de Regiones De Interés (RDI).

Al igual que su semejante, el modelo de bordes ignora mucha de la información contenida en el resto de la muestra, y la región observada debe ser equivalente para las distintas clases. En caso del Euro, que su valor nominal figura en sus bordes, puede ser una posibilidad potente para el reconocimiento del billete, ya que simplifica mucho la imagen y permite la aplicación directa de reconocimiento del patrón de números.

5.4. Aplicación específica de PCA

En el capítulo 4.4, se ha discutido de forma deliberada el Análisis de Componentes Principales (PCA) para evitar extenderse innecesariamente en esta sección. Lo que se intenta defender aquí es la transformación de la imagen en una “señal” tratable. Esto se consigue convirtiendo la imagen a una forma vectorial. Así pues, la transformada de una imagen a señal se realizará

organizando los píxeles de la muestra en una secuencia ordenada de una dimensión (forma vectorial).

Como en las secciones anteriores, el modo en que la imagen se transforme en señal influye en el código y los parámetros del reconocimiento de las muestras. A continuación, se presentan distintas opciones que dan lugar a diferentes resultados.

5.4.1. Columnas/Filas ordenadas

Se presenta la propuesta de ordenar las filas o columnas en forma de vector una de tras de otra. Cuando la organización de la imagen se realiza de este modo, los valores de la señal se mantienen en el rango de valores de la imagen, y el número de muestras de las que se compone la señal se corresponde al número de píxeles que conformaban su imagen. Como resultado, se obtienen componentes largas (de gran número) de amplitud acotada entre 0 y 255.

5.4.2. Columnas/Filas acumuladas

Se presenta como propuesta acumular el valor de los píxeles de una misma columna, o fila, en una muestra de la señal. El número de muestra se corresponderá al número de columna o fila acumulada.

Cuando la organización de la imagen se realiza de este modo, acotar el rango de valores de las muestras de la señal no es posible. Sin embargo la señal es mucho más corta; es decir, se compone de un número mucho más reducido de muestras. Como resultado, se obtienen componentes cortas de gran amplitud.

Se podría defender que debido al reducido número de muestras que forman la señal, el procedimiento descrito en el apartado 5.4.3 se vuelve innecesario. Si se compara el método acumulativo (apartado actual, 5.4.2) con el método de ordenación (apartado anterior, 5.4.1) se deduce fácilmente que el acumulativo utiliza un menor número de bytes.

5.4.3. Reducción de la muestra

En el apartado 5.2.4 se ha mencionado el inconveniente de las diferentes dimensiones para las distintas clases de muestras. También se menciona la decisión de reducir las muestras a una dimensión común, preferentemente la menor. Aún así, la dimensión de las muestras sigue siendo grande, y se plantea la opción de reducir aún más esta.

Se considera un “Selector” del método debido a que su valor influye en los resultados de la clasificación además de la carga de procesado; se ha observado que a medida que se reducen las dimensiones de la referencia y la muestra la fiabilidad del reconocimiento aumenta ligeramente hasta llegar a cierto tamaño. En este valor, la fiabilidad del reconocimiento adquiere comportamiento parabólico y se reduce. Se interpreta este comportamiento como que a medida que el tamaño se reduce llega un punto en que las distintas clases se asemejan demasiado (los detalles que las diferencian dejan de percibirse) y el reconocimiento pierde fiabilidad.

5.5. Estudio de la fiabilidad y eficiencia: *Selectors*

Los denominados en este proyecto “Selectores” son medidas y valores (normalmente numéricos) que influyen con cierta importancia en los resultados del reconocimiento de

muestras y señales. Es decir, la eficiencia del modelo y método dependen de estos “Selectores”.

5.5.1. Dimensión de la muestra referencia

En este punto se recompila y complementa el trato de la dimensión de la señal y de la muestra. Como se ha mencionado con anterioridad, las imágenes de billetes no comparten entre ellas mismas dimensiones, siendo este un punto necesario para la aplicación de PCA. También, como se ha descrito, uno de los objetivos es reducir la carga de procesado, y una de las posibilidades es la reducción de la señal que luego ha de descomponerse en CPs.

Se considera un “Selector” del método debido a que su valor influye en los resultados de la clasificación además de la carga de procesado; se ha observado que a medida que se reducen las dimensiones de la referencia y la muestra la fiabilidad del reconocimiento aumenta ligeramente hasta llegar a cierto tamaño. En este valor, la fiabilidad del reconocimiento adquiere comportamiento parabólico y se reduce. Se interpreta este comportamiento como que a medida que el tamaño se reduce llega un punto en que las distintas clases se asemejan demasiado (los detalles que las diferencian dejan de percibirse) y el reconocimiento pierde fiabilidad.

5.5.2. Discriminación de componentes

La varianza es un “Selector” que se relaciona profundamente con la elección de PCA como método de clasificación. Se recuerda que las señales de referencia se descomponen en varias PCs y que cada una de estas tiene una desviación típica (o varianza) asociada para el conjunto de señales. La suma de estas da como resultado un valor de varianza del conjunto.

Si se utilizan las PCs para discriminar aquellas componentes con menor efecto en el conjunto de señales la varianza total del conjunto también disminuye. Para la implementación, se utilizará el planteamiento inverso: “si se especifica una varianza máxima para el conjunto, las componentes de menor peso serán discriminadas”.

Pensar en este procedimiento como un cambio continuo es una equivocación. Aunque la varianza del conjunto es una variable continua, el número de componentes asociado no, y hasta que la varianza no alcanza cierto valor (que no es predecible hasta aplicar PCA sobre el conjunto) el número de componentes incluidas no cambia. El efecto de recudir severamente la discriminación de componentes hace que el proceso de reconocimiento falle; se ha comprobado que para muchas de las muestras la reducción de las componentes a 0.8 de variabilidad, causa que se confundan con otras clases.

La varianza del conjunto no es necesaria normalizarla para implementarla como “Selector”, sin embargo se normaliza para poder expresar la varianza aceptada del conjunto como una proporción sobre 1. Esta valoración ayuda a la comprensión cuando se están realizando pruebas y cambios en otros selectores y en el propio modelo. Es, en esencia, una mejora práctica.

5.5.3. Umbral de confianza y tendencia (*Threshold and bias*)

Se refiere el “umbral de confianza” al indicador (normalmente una variable numérica) que determina, en este caso, cuando la similitud entre una muestra y un modelo ideal de una clase permite que la muestra se considere de la clase en cuestión.

En la implementación del software, este umbral es un valor numérico entre 0 y 1 (debido a la normalización de la correlación). Se normaliza la correlación para poder valorar en caso de conflicto entre dos clases distintas. Se recuerda que debido al modelo de PCA, las componentes que forman las señales de referencia son comunes en forma para todas las clases, y que la matriz W_{PCA} , es la que contiene los factores que se utilizan para la reconstrucción. Así pues, puede existir un caso en que la correlación de dos señales sobrepase el umbral de confianza, momento en que habrá que decidir entre ambos. Normalmente, la muestra se identificara con el valor de correlación mayor.

El umbral de confianza es ante todo un criterio para discriminar; implementa un criterio que se puede entender como una exigencia de valor para la correlación. Si se considera no incorporar esta exigencia, se puede evaluar que correlación es la más alta (una vez han sido normalizadas) y clasificar en función de esta. Esta simplificación puede causar la aceptación de muestras que no pueden caracterizarse; muestras que coinciden con las clases existentes pero no pertenecen a ellas. Una reducción importante de la varianza también puede causar este efecto.

La tabla de la Figura 5-9, se quiere mostrar la similitud existente entre las distintas clases cuando la señal de referencia y la muestra son iguales. En el caso planteado se refleja el hecho que, pese a ser claramente imágenes distintas, debido al método de PCA aplicado, la correlación tiene un valor muy alto con otras clases. Con esto quiere defenderse la importancia de una buena selección del umbral de confianza.

		Muestra de entrada										
		5#00	10#00	20#00	50#00	100#00	200#00	500#00	5N#00	10N#00	20N#00	50#00N
Señal de referencia	5	0,9974	0,9003	0,9131	0,8528	0,7318	0,8486	0,8303	0,7116	0,7031	0,6477	0,6586
	10	0,9149	0,9978	0,9314	0,8345	0,6666	0,845	0,8346	0,7149	0,6901	0,6332	0,6277
	20	0,9119	0,9151	0,9978	0,8885	0,736	0,8716	0,8721	0,7274	0,7274	0,699	0,681
	50	0,8516	0,8201	0,8883	0,9963	0,6651	0,8886	0,9282	0,6694	0,6899	0,6906	0,6664
	100	0,7202	0,6458	0,7257	0,6565	0,997	0,6695	0,6795	0,5598	0,588	0,5656	0,5846
	200	0,8557	0,8384	0,8799	0,8972	0,6852	0,9974	0,9272	0,5857	0,5937	0,616	0,6009
	500	0,8168	0,8082	0,8592	0,9146	0,6788	0,9048	0,9978	0,6297	0,6721	0,6946	0,664
	5N	0,7199	0,7118	0,7371	0,6783	0,5759	0,5878	0,6475	0,9991	0,903	0,8421	0,8092
	10N	0,7035	0,6796	0,7291	0,6914	0,5979	0,5895	0,6838	0,893	0,9987	0,8678	0,858
	20N	0,6587	0,634	0,7122	0,7035	0,5852	0,6217	0,7183	0,8464	0,8818	0,9981	0,9332
	50N	0,6585	0,6177	0,6821	0,6675	0,5949	0,5961	0,6749	0,7997	0,8575	0,9175	0,9994

Figura 5-9, Resultados del proceso de reconocimiento sin reducción de datos por selectores.

La tendencia es un indicador para modelos adaptativos. La tendencia indica como progresa el modelo de referencia tras clasificar las nuevas muestras, ya sea convergiendo los posibles valores de clase en un patrón poco variable, o divergiendo de este.

La tendencia también puede aplicarse al umbral de confianza; si el umbral es un indicador adaptativo, se espera que varíe y ajuste con el tiempo. Su adaptación puede causar el fallo del sistema al completo, volviéndose demasiado permisivo o demasiado crítico.

6. Conclusiones y trabajo a futuro

6.1. Conclusiones

Los resultados obtenidos durante las pruebas con muestras concluyen que el método planteado en este documento (reconocimiento mediante PCA) tiene un gran potencial para reconocer billetes con un escaneo de su imagen. Así mismo, también se ha concluido que los parámetros como la variabilidad del conjunto de PCs y el tamaño de las muestras analizadas debe tenerse muy en cuenta cuando intenta reducirse la carga de procesado; si su influencia se ignora, el método pierde eficacia de forma drástica.

6.2. Trabajo a futuro

A lo largo del documento, y sobre todo en el capítulo 5 (Implementación de *software*), se han descrito métodos que, por diversas razones, no han sido implementados o no se han estudiado a fondo. Como trabajo futuro, podría estudiarse su implementación en el software, ya fuere substituyendo partes actuales o como refuerzo de estas.

El estudio sobre las Regiones de Interés se ha ignorado por su complejidad, al igual que la utilidad de imágenes infrarrojas o de otras radiaciones para la detección de falsificaciones, números de serie, etc.

También, la mejora del procesado de imagen es un trabajo constante; lamentablemente muchas de las imágenes escaneadas que se proveyeron fueron rechazadas para utilizarse en código debido a que el procesado de imagen generalizado no daba un resultado adecuado con dichos escaneos.

Bibliografía

Net y redes

- ❖ www.riksbank.se/en/The-Riksbank/History/Money-and-power-the-history-of-Sveriges-Riksbank/Stockholms-Banco
- ❖ <http://homepages.inf.ed.ac.uk/rbf/HIPR2/fourier.htm>

Artículos

- ❖ Bala et al., *International Journal of Emerging Research in Management & Technology* ISSN: 2278-9359 (Volume-3, Issue-7)
- ❖ Masato Aoba, Tetsuo Kikuchi, Yoshiyasu Takefuji, "Euro banknote recognition system using a three layer perceptron and RBF networks", *IPSI Transaction on Mathematical Modeling and Its Application*, Vol 44, No. SIG 7 (TOM 8), May 2003, Pp. 99-109.
- ❖ Ali Ahmadi, Sigeru Omatu, Toshihisa Kosaka, "A Reliable Method for Recognition of Paper Currency by Approach to Local PCA) used local principal component analysis PCA", in *IEEE proceedings of the International Joint Conference on Neural Networks*, 20-24 July 2003, Vol. 2, Pp. 1258 – 1262.
- ❖ D. A. K. S. Gunaratna, N. D. Kodikara and H. L. Premaratne, "ANN Based Currency Recognition System using Compressed Gray Scale and Application for Sri Lankan Currency Notes-SLCRec", in *proceedings of world academy of science, engineering and technology*, Nov 2008, vol. 35, ISSN 2070-3740, Pp. 235-240.
- ❖ Kalyan Kumar Debnath, Sultan Uddin Ahmed, Md. Shahjahan, "A Paper Currency Recognition System Using Negatively Correlated Neural Network Ensemble", *Journal Of Multimedia*, December 2010, Vol. 5, No. 6, Pp. 560-567.
- ❖ Junfang Guo, Yanyun Zhao, Anni Cai, "A reliable method for paper currency recognition based on LBP" in *proceeding of 2nd IEEE International Conference on Network Infrastructure and Digital Content*, 24-26 Sept. 2010, Pp. 359 – 363.
- ❖ Hai-dong Wang, Leye Gu, Linping Du, "A paper currency number recognition based on fast Adaboost training algorithm", in *IEEE proceedings of International Conference on Multimedia Technology (ICMT)*, Pp. 4772 – 4775.
- ❖ Chetan B. V., Dr. P. A. Vijaya, "A Robust Side Invariant Technique of Indian Paper Currency Recognition", *International Journal of Engineering Research & Technology (IJERT)*, May – 2012, Vol. 1 Issue 3, Pp. 1-7.

Libros

- ❖ Sing-Tze Bow, *Pattern Recognition and Image Preprocessing* Second Edition, Revised and expanded ISBN:0-8247-0659-5
- ❖ *Handbook Of Pattern Recognition & Computer Vision* (3RD Edition), ISBN 981-256-105-6
- ❖ Stone, James V., *Independent component analysis : a tutorial introduction*, ISBN 0-262-69315-1

Glosario y conceptos

- *Banknote*: Billeto.
- Blind Signal Separation (BSS): Separación Ciega de Señales. Véase en la página 20.
- *Boundary conditions*: Condición de contorno.
- *Bounded signals*: señales discontinuas con una o varias tendencias a $\pm \infty$.
- Complejidad temporal (informática): es la medida de la cantidad de recursos utilizada por un algoritmo.
- *Currency*: Moneda.
- *Data Set*: Grupo de datos.
- Dependencia lineal: Un conjunto de variables aleatorias son linealmente dependientes cuando al menos una de las señales del conjunto puede expresarse como combinación lineal de una o varias del resto de señales del conjunto.
- Entropía (estadística, información): Es una medida de la incertidumbre de una variable o evento aleatorio.
- Fiabilidad: Medida de cuan fiable es el método de reconocimiento.
- Función de densidad de probabilidad (*pdf*): función que representa la probabilidad de que una variable aleatoria tome cierto valor.
- Función monótona creciente: una función se dice que es monótona creciente cuando la expresión de su derivada tiene solamente valores positivos para su recorrido.
- Gradiente (matemáticas): Vector que es perpendicular a la curva de nivel en el punto en cuestión.
- Independencia estadística: Dos señales aleatorias son estadísticamente independientes cuando la señal $\langle a \rangle$ no puede expresarse como una composición de ningún orden de la señal $\langle b \rangle$.
- Independencia lineal: Un conjunto de variables aleatorias son linealmente independientes cuando ninguna de las variables puede expresarse como una composición lineal del resto.
- *Mixing matrix*: Matriz mezcladora.
- *Pattern*: Patrón. Puede describirse como una descripción cuantitativa o estructural de un objeto o entidad. Preprocesado de imagen: Etapa previa al procesamiento de imagen que realiza cambios con el objetivo de mejorar o facilitar los resultados de su etapa posterior (procesado de imagen).
- *Signal mixture*: Señal mezcla. Ir a página 20.
- *Source signal*: Señal fuente. Ir a página 20.
- *String* (programación): Cadena de caracteres.
- *Template image*: “Imagen plantilla”.
- *Unmixing matrix*: Matriz recuperadora.

Índice de figuras

Figura 3-1, Recreación de un cartel holográfico "Tranquillity's STOP"	7
Figura 3-2, Detalle de los pixeles del cartel.	8
Figura 3-3, Fotografía de un edificio acristalado.	9
Figura 3-4, Mejora aplicando un escalado lineal.	9
Figura 3-5, Diferentes aplicaciones de transformada de nivel.	10
Figura 3-6, Relación entre el histograma de una variable y grafica densidad de probabilidad.	11
Figura 3-7, (a) Imagen artificial rectangular y (b) su espectro.	12
Figura 3-8, Esquema de las etapas de la imagen según el dominio utilizado.	13
Figura 3-9, (a) Imagen desplazada y (b) su espectro tras el desplazamiento.	14
Figura 3-10, (a) Imagen artificial rotada y (b) su espectro tras rotar.	15
Figura 3-11, Efecto de la reducción del espectro: (a) Imagen original, (b) Interpretación de su espectro, (c) Imagen restaurada de un 5% del espectro, (d) Imagen restaurada de un 10% de su espectro.	16
Figura 3-12, Efecto de la aplicación del operador de Sobel; (a) imagen original, (b) tras aplicar el operador horizontal, (c) tras aplicar el operador vertical.	18
Figura 3-13, Escaneo de billete tras sustraer el trasfondo.	18
Figura 4-1, A la izquierda, distribución de dos variables aleatorias independientes. A la derecha, distribución de dos variables con dependencia.	20
Figura 4-2, Señal y forma; Izquierda, señal super-gaussiana; Centro, señal dientes de sierra; Derecha, señal gaussiana.	21
Figura 4-3, Esquema sobre la interpretación de la función g	26
Figura 5-1, Esquema de la organización de las imágenes escaneadas.	33
Figura 5-2, Imágenes originales de los ejemplos: (a) 50R#013, (b) 10NAI#032.	34
Figura 5-3, Diagrama de bloques que representa las etapas del <i>software</i>	35
Figura 5-4, Imagen sin modificar de la muestra 100A#010.	36
Figura 5-5, Selección de traza y su forma.	37
Figura 5-6, (a) imagen binaria donde se puede diferenciar el objeto y (b) contorno del objeto tras aplicar el operador de Sobel.	38
Figura 5-7, Tabla con las dimensiones medias de las imágenes en pixeles y su desviación estándar. ...	39
Figura 5-8, (a) Imagen idealizada (b) muestra sin modificar.	40
Figura 5-9, Resultados del proceso de reconocimiento sin reducción de datos por selectores.	43

Apéndices

Apéndice P: Códigos de programa

getIfromdata()

```
function I=getIfromdata(n1,n2,n3)
% <getIfromdata> carga del directorio especificado los archivos útiles para
% el software. Utilizado un código de tres números se especifica una única
% imagen.

directorio='';
face={'A','AI','R','RI'};
cara=face{n1};
value=[5,10,20,50,100,200,500,5,10,20,50];
valor=value(n2);
year=[2013,2014,2015,2017];

if n3<10
    num=['00',num2str(n3)];
else
    num=['0',num2str(n3)];
end
if n2<8 %Modelo standar.
    I=imread([directorio,'euro/',num2str(valor),cara,'/Cis1Phase6_',...
        num2str(valor),cara,'_',num,'.bmp']);
else %Modelo nuevo.
    I=imread([directorio,'euro_nuevo/',num2str(valor),'E',...
        num2str(year(n2-7)),cara,'/Cis1Phase6_',num2str(valor),'E',...
        num2str(year(n2-7)),cara,'_',num,'.bmp']);
end

end
```

imageProcess()

```
% <imageProcess>. Sintetiza todo el procesado de imagen.
% I y A son dos estados de la misma imagen. I es la imagen sin procesar.
% A es la imagen I tras su procesado.

function A=imageProcess(I)
[y]=preparaEnmarcado(I);
%Los valores <m> y <b> son la pendiente y ordenada de las rectas que
%definen el contorno.
[m_hs,m_hi,m_v1,m_v2,b_hs,b_hi,b_v1,b_v2]=enmarca_billete_v1(y);
%Los puntos <p> son los puntos de intersección de las rectas.
[p11,p12,p21,p22]=calculaEsquinas(m_hs,m_hi,m_v1,m_v2,b_hs,b_hi,b_v1,b_v2);

%Rotación. Se mantiene el tamaño original de la imagen.
angled=rad2deg(atan(m_hs)); %Angulo en radianes.
angler=atan(m_hs); %Angulo en grados.
I6=imrotate(I,angled,'nearest','crop');
mpun=round([p11;p12;p21;p22]);
mfc=rota_puntosV2(mpun,I,angler);
%Crop image.
A=I6(min(mfc(:,2)):max(mfc(:,2)),min(mfc(:,1)):max(mfc(:,1))));

end
```

buildpatternV3()

% <build_patternV3> construye modelos de las imágenes de billetes de euro.

```
function build_patternV3()
N1=4;N2=11;N3=49; %Limits.
Nclass=22;
Dim=[100,200]; %Selector
Xsc=zeros(Nclass,Dim(1)*Dim(2));

for n1=[1,3] %Solo caras 'A' y 'R'
    for n2=1:N2
        S=zeros(Dim(1),Dim(2));
        n=0;
        for n3=1:10 %N3
            n=n+1;
            I=getIfromdata(n1,n2,n3);
            I=imrotate(I,90);
            I=imageProcess(I);
            I_reducida=imresize(uint8(I),Dim); %Reducción.
            S=S+double(I_reducida);
        end
        S=S/n;
        %Transformada a señal.
        if n1==1
            Xsc(n2,:)=reshape(S,Dim(1)*Dim(2),[]);
        elseif n1==3
            Xsc(n2+N2,:)=reshape(S,Dim(1)*Dim(2),[]);
        end
    end
end
%PCA
[Xsc,m]=remmean(Xsc);
[weights,~,compvars,~,~,~,~]=runica(Xsc,'verbose','off','pca',N2*2);
Xpc=weights*Xsc; %PCs
var_acuN=cumsum(compvars/sum(compvars));
%Norma de PCs
W=inv(weights);
XpcN=zeros(size(Xpc));
WN=zeros(size(W));
Xpc_mod=sqrt(diag(W*W'));
W_mod=sqrt(diag(Xpc*Xpc'));
for i=1:size(W,2)
    WN(i,:)=W(i,:)/Xpc_mod(i);
    XpcN(i,:)=Xpc(i,:)/W_mod(i);
end
%Guardado del pattern.
directorio='';
saveroute=[directorio,'patterns/','pattern-master3'];
save(saveroute,'Xpc','W','XpcN','WN','var_acuN');
%Variables guardadas para el modelo.

end
```

rec_patternV3()

% Reconoce billetes con ayuda del modelo de patrones V3.

```
function rec=rec_patternV3(n1,n2,n3,num)
% Carga los datos.
load patterns/pattern-master3 %directorio

%Recupera la dimensión para prop=2.
Dim=[0 0];
Dim(1)=sqrt(length(Xpc(1,:))/2);
Dim(2)=Dim(1)*2;
Xsc=zeros(1,length(Xpc(1,:)));

%Procesado.
I=getIfromdata(n1,n2,n3);
I=imrotate(I,90);
I=imageProcess(I);
%Reducción a común dimensión y modelado a señal.
I_reducida=imresize(uint8(I),Dim);
Xsc(1,:)=double(reshape(I_reducida,Dim(1)*Dim(2),[]));
%Reducción de componentes.
if num<1
    aux=find(var_acuN>=num);
    n_comp=aux(1);
    clear aux
else
    n_comp=num;
end
%Extracción del vector de pesos.
ss=Xsc*XpcN';
ssN=ss/sqrt(ss*ss'); %Normalización del vector pesos.
%Correlación. <r> es el vector de correlaciones
r=WN(:,1:n_comp)*ssN(:,1:n_comp)';

%Identificación. String de salida.
A=r/max(r);
aux=find(A==1);
if (aux<12)
    cara='A';
else
    cara='R';
    aux=aux-7;
end
value=[5,10,20,50,100,200,500,5,10,20,50];
id=num2str(value(aux));
if aux>7
    id=[id,'N']; %Información de nuevo modelo.
end
id=[id,cara];
%Struct.
rec.id=id;
rec.nid=find(A==1);
rec.corr=r;
rec.corrn=r(A==1);
rec.Dim=Dim;
rec.comp=n_comp;

end
```


Apéndice F: Códigos de figuras del capítulo 5

Figura 5-2

```
% Figure 5-2. Imágenes sin tratar de los ejemplos: 50R#013, 10NAI#032.
I1=getIfromdata(3,4,13); %50R#013
I1=imrotate(I1,90);
I2=getIfromdata(2,9,32); %10NAI#032
I2=imrotate(I2,-90);

figure('Color', [1 1 1])
subplot(121), imshow(I1), title('(a) 50R#013')
subplot(122), imshow(I2), title('(b) 10NAI#032')
```

Figura 5-4

```
% Figure 5-4. Imagen sin modificar de la muestra 100A#010.
I1=getIfromdata(1,5,10); %100A#010
figure('Color', [1 1 1])
imshow(I1), title('100A#010')
```

Figura 5-5

```
% Figure 5-5. Selección de traza y su forma.
I1=getIfromdata(1,5,10); %100A#010
figure('Color', [1 1 1])
imshow(I1), title('100A#010')
trace=I1(15,:);
hold on, line([0,size(I1,2)],[15,15]), hold off
figure('Color', [1 1 1])
plot(trace), axis image
```

Figura 5-6

```
% Figure 5-6. (a) imagen binaria donde se puede diferenciar el objeto y (b)
% contorno del objeto tras aplicar el operador de Sobel.
% Se ha utilizado código original de <preparaEnmarcado> para su contrucción.
```

```
I1=getIfromdata(1,5,10); %100A#010
figure('Color', [1 1 1])
    imshow(I1), title('100A#010')
nv=255-I1;
traza=nv(15,:);
fondo=repmat(traza,size(nv,1),1);
dif=nv-fondo;
ct=contraste(dif,20,200,4,170);
bw=im2bw(ct,22/256);
bw=imrotate(bw,90);
    subplot(211),imshow(bw), title('(a)')
S=[-1 -1 0;-1 0 1;0 1 1];
perfil=logical(abs(conv2(S,double(bw))));
    subplot(212),imshow(perfil), title('(b)')
```

Figura 5-9

```
% Figure 5-9. Resultados del proceso de reconocimiento sin reducción de
% datos por selectores.
reducirS=true; fil=13; col=18;
```

```

reducirC=true; varN=double(.999); %Varianza permitida.

res=zeros(11);
%-----%
face={'A','AI','R','RI'};
value=[5,10,20,50,100,200,500,5,10,20,50];
year=[2013,2014,2015,2017];
N1=size(face,2);
N2=length(value);

%Seccion 1: Build the pattern.
n1=1; n3=0;
Xs=zeros(N2,fil*col);
for n2=1:N2
    I=getIfromdata(n1,n2,n3);
    I=imageProcess(I);
    if reducirS
        [h,w]=size(I);
        if h>w %Pasa de vertical a Horizontal.
            I=imrotate(I,90);
        end
        I=imresize(I,[fil,col]);
    end
    Xs(n2,:)=reshape(I,fil*col,[]);
    ... %¿que más?
end
Xs=remmean(Xs);
[weights,spheres,compvars]=runica(Xs,'verbose','off','pca',N2);
Xp=weights*Xs;
W=inv(weights);
%1° Normalizar, 2° reducir.
mod_Xp=sqrt(diag(Xp*Xp'));
mod_W=sqrt(diag(W*W'));
XpN=zeros(size(Xp));
WN=zeros(size(W));
for i=1:N2
    WN(i,:)=W(i,:)/mod_W(i);
    XpN(i,:)=Xp(i,:)/mod_Xp(i);
end
if reducirC
    varN_acu=cumsum(compvars/sum(compvars));
    num_c=find(varN_acu>=varN);
    num_comp=num_c(1);
    %var_ac=vracu(num_comp); %Info complementaria.
    W_r=WN(:,1:num_comp);
    Xp_r=XpN(1:num_comp,:);
end

%Seccion 2: Reconocimiento.
id=zeros(1,N2);
for n2=1:N2
    ss=Xs(n2,:)*XpN';
    ssN=ss/sqrt(ss*ss');
    similitud=WN*ssN';
    V=similitud/sum(similitud); %V es la similitud nominal.
    ...
    res(:,n2)=similitud;
    [~,aux]=max(similitud);
    id(n2)=value(aux);
end

```